Important Notice

UNIVERSITY OF CALGARY

WAVEFORM INVERSION FOR AREAS WITH COMPLEX NEAR SURFACE

by

Hussain I. Hammad

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF GEOSCIENCE

CALGARY, ALBERTA

February, 2010

UNIVERSITY OF CALGARY

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled "WAVEFORM INVERSION FOR AREAS WITH COMPLEX NEAR SURFACE" submitted by Hussain I. Hammad in partial fulfillment of the requirements for the degree of MASTER OF SCIENCE.

Supervisor, Dr. Gary F. Margrave
Department of Geoscience

Dr. Larry R. Lines
Department of Geoscience

Dr. David W. Hobill
Department of Physics and Astronomy

Date

Abstract

Seismic waveform inversion is a method with very ambitious goals. It attempts to obtain Earth's model parameters from seismic data in one encompassing and comprehensive process. This thesis, however, is concerned with acoustic waveform inversion in the frequency-domain with emphasis on areas with complex near surface.

We give an overview of seismic inversion and the method on which waveform inversion is based, seismic migration. We review waveform inversion from a theoretical point of view and attempt to coherently put together the different theories and conclusions reached by some authors in the literature. We derive the formula of the gradient of the misfit function for waveform inversion and, in one derivation, we explicitly invoke the well known linearization, the Born approximation, and therefore we show how it is related to seismic migration. We, also, compare and contrast waveform inversion with seismic migration, in general, Kirchhoff migration and true amplitude inversion, in particular; in addition to traveltime tomography.

Since the near surface is highly heterogeneous, we examine the effect of heterogeneity on waveform inversion and whether it can perform in an acceptable manner to relatively high frequency, almost 25 Hz. We then apply waveform inversion to a realistic model representative of challenging near surface environments, like those in the Middle East. Waveform inversion was able to give very high resolution models that can resolve the issues associated with long-wavelength statics. We then examine if waveform inversion can indeed resolve the challenging problem of first arrival shingling. We show that it can indeed resolve this long standing problem.

Acknowledgments

First and foremost, I thank the Lord, the God of the Universe, for his grace, bounty and endless blessings. I thank the many people who made finishing this thesis possible. I thank, especially, Dr. Gary Margrave, my supervisor, for his support, guidance and enlightening discussions. I definitely learned a lot from his thoroughness and rigor. His wonderful lectures, particularly those on inverse scattering theory, have definitely influenced this thesis, especially chapter 2. I thank Dr. Larry Lines, who may not be aware, that his early help was absolutely instrumental in getting me started on this project. I thank Dr. Rob Stewart for his early enthusiasm, encouragement and support at the early stages of this project. Dr. Pat Daley is thanked for his review and for the post-midnight discussions and heated debates. Dr. Jo Wong and Dr. Peter Manning are also thanked for their review. Dr. Rolf Mayer and Kevin Hall are thanked for their technical support. I thank all CREWES staff, faculty, my colleagues and friends for their help and companionship throughout this journey. Special thanks go to Dr. Gerhard Pratt, who is now at the University of Western Ontario, for his wonderful course and codes and for helping me get started on this subject. I thank Saudi Aramco for supporting my studies. Many thanks go to Riaydh Al-Saad, who introduced me to the problem of shingling and Dr. Yi Lu for the discussion and the model of shingling. Dr. Saleh Al-Saleh and Shelton Hubble are also thanked for their help in providing the near surface model. Last but not least, nothing could have been achieved without the endless support, love, prayer and supplication of my beloved family.

Dedication

To my mother

and for the memory of my father.

Table of Contents

List of Figures

Chapter 1

Introduction

Seismic waveform inversion attempts to achieve a very ambitious goal, namely obtain Earth's model parameters from seismic data in one encompassing and comprehensive process. In this chapter we conceptually review waveform inversion and the method on which it is based, seismic migration. We summarize the thesis and we state the objectives as well as the tools used.

1.1 Introduction

Seismology is concerned with inferring the interior of the Earth by recording and studying its mechanical vibrations. The vibrations are either caused by natural sources such as earthquakes or man-made sources such as explosions (Aki and Richards, 2002, p. 1). Exploration seismology, which is our subject of interest, is concerned specifically with the top few kilometers of the Earth crust for the purpose of exploration and development of natural resources such as petroleum (see Yilmaz, 2001; Sheriff and Geldart, 1995).

1.2 Seismic data acquisition

Before we dive into the details, we attempt first to describe how the seismic data, we are concerned with, are acquired. A more thorough treatment of the subject of seismic acquisition is found in many books such as Yilmaz (2001); Sheriff and Geldart (1995); Liner (2004) and Vermeer (2002). We are mainly concerned, in this thesis, with 2D surface seismic surveys. Those 2D seismic surveys are acquired on the surface of the Earth. The assumption embedded in using such a geometry is that the variations in the

geology perpendicular to the profile are weak (Tarantola, 1984b). The receivers used in the surveys can be one component receivers, three component or even four component receivers. One component receivers measure the vertical (to the acquisition surface) particle velocity or pressure in the case of a hydrophone. Adding two horizontal receivers, we can measure particle velocity in the two planes perpendicular to the plane of the vertical component, as shown in Figure 1.1. In this thesis, however, we are mainly concerned with the vertical component.

The source used can be a dynamite explosion or a truck with a vibrating plate, vibroseis truck, on land surveys or an air gun in marine surveys. Those sources emit waves that travel into the subsurface and scatter in all directions. The receivers on the surface receive the energy that return to the surface, where they record the wave field pressure, particle velocity or acceleration depending on the type of receiver.

1.3 Inversion

Once the seismic data are acquired, we can then use them to infer the properties of the Earth's interior. The process of the inversion allows us to map the data from their space to the model space which is, in our case, the space of the Earth's parameters of interest as shown in Figure 1.2. Inversion is essentially composed of two steps: an estimation step and an appraisal step (Snieder, 1998). The estimation process involves estimating the model parameters, whereas the appraisal process involves evaluating how close the estimated model is to the true model. As Snieder (1998) mentions, the appraisal process is well understood for linear problems but that is not the case for the more difficult non-linear problems, such as waveform inversion.

As Weglein et al. (2009) classify it, inversion is divided into two types: direct and indirect. Direct inversion outputs the model directly from the data: data in, model out.

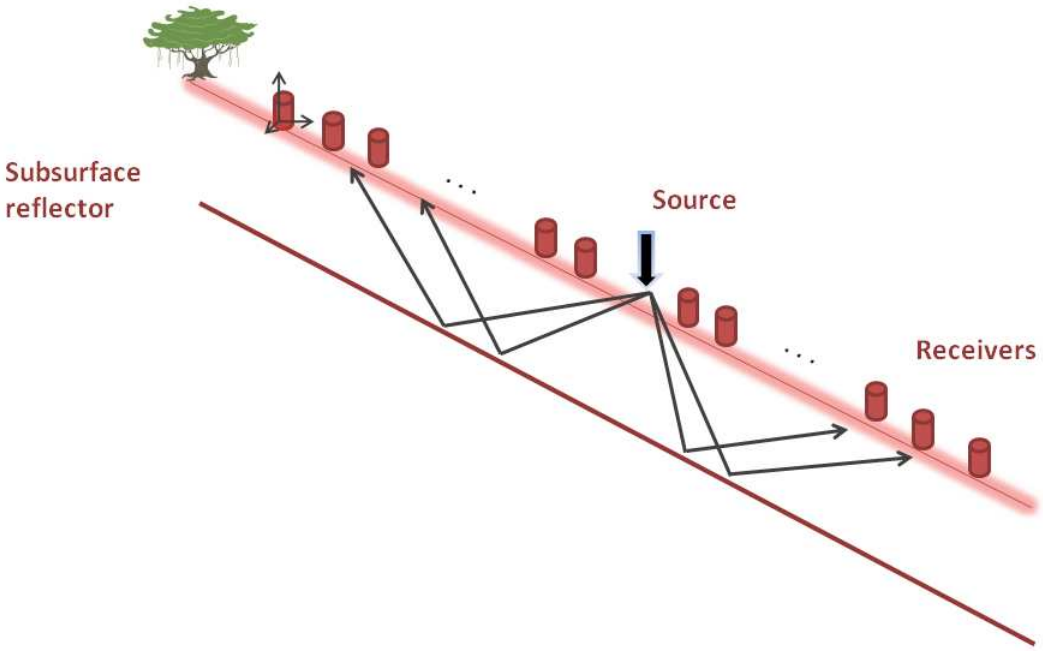


Figure 1.1: Schematic of 2D seismic acquisition survey.

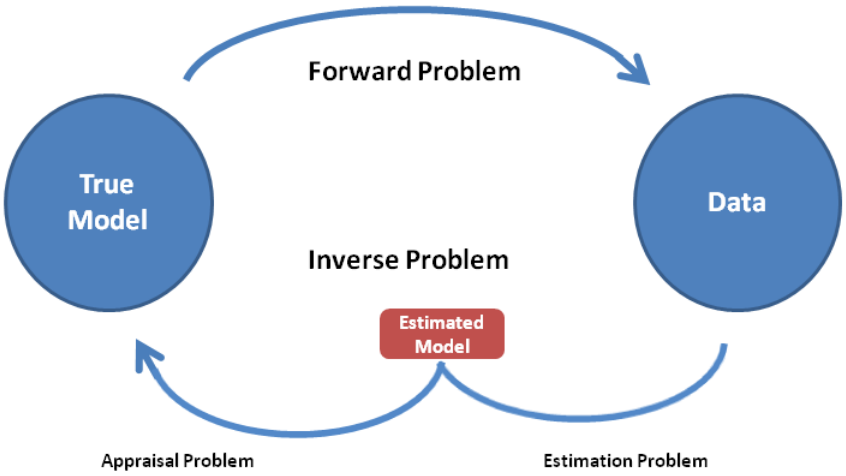


Figure 1.2: A schematic that shows the relation between forward modeling, inversion and the member processes of inversion, the estimation and the appraisal problem. Figure is adapted from Snieder (1998).

Indirect inversion, on the other hand, outputs a model through an intermediate step such as optimization based inversion, which attempts to minimize an objective function that involves the observed data and the modeled data. An example of direct inversion is the inverse scattering series (e.g. Weglein et al., 2003, 2009) and an example of indirect inversion is waveform inversion.

We describe here the general flow of inversion. This is thoroughly discussed in any standard book on geophysical inversion such as Tarantola (2005); Parker (1994); Menke (1989); Aster et al. (2005) and Claerbout (1992). The general process of optimization-based inversion, that uses local methods, is as follows :

1. Start with a background model. Set the background model to be the current model.
2. Generate synthetic data using the current model.
3. Subtract the synthetic data from the observed data to create the data residuals.
4. Update the model in such a way that minimizes the data residuals.
5. Repeat 2-5 until a certain criterion is met.

Inverse problems are generally very difficult. Aster et al. (2005, p. 11-12) give some reasons why they are so. They mention that their difficulty stems from three issues: existence, uniqueness and instability. Inverse problems may not have a solution because the physics of the forward problem is too approximate or due to the noise in the data. There might be infinitely many models that can fit the data. Inverse problems might be ill-posed or ill-conditioned, where the solution is very sensitive to small changes in the data.

Waveform inversion, in particular, is concerned with finding the models whose waveforms best match the observed waveforms. It is based on seismic migration, so we will first discus seismic migration.

1.4 Seismic migration

Sheriff (2002) in his dictionary defines seismic migration as an "inversion operation involving rearrangement of seismic information elements so that reflections and diffractions are plotted at their true locations." So the goal of migration is to to produce an accurate image of the geometry or the structure of the subsurface. It also aims at producing an image that estimates the band limited reflectivity of the Earth. The reflectivity as Sheriff (2002) defines it is "the ratio of the amplitude of the displacement of the a reflected wave to that of the incident wave." The commonly used seismic migration algorithms are based on some assumptions. Here we list the general assumptions and more detailed discussion can be found in Claerbout (1985); Berkhout (1985); Yilmaz (2001); Stolt and Benson (1986) and Biondi (2006). First, it is often based on the acoustic approximation due mainly to computational cost. Second, it assumes that the velocity model is known. Third, it is based on the first order Born approximation. In other words, it assumes that the data have only single scattering as opposed to multiple scattering, as we will discuss in detail in Chapter 2. Fourth, it assumes that the incident wavelet is known. Additional assumptions might also be introduced by the specific migration method such as phase shift methods that assume mild lateral velocity variations (e.g. Gazdag, 1978; Gazdag and Sguazzero, 1984).

Seismic migration methods can be classified into two generic categories: wavefield extrapolation methods and Kirchhoff or integral methods (Larner and Hatton, 1990; Bevc and Biondi, 2005).

Wavefield extrapolation methods can be divided into two methods: those based on the one-way wave equation and those based on the two-way wave equation (Sava and Hill, 2009). Those based on the two-way wave equation are also referred to as reverse time migration. For the one-way methods, the wavefield extrapolation is carried out in

depth. The one-way wave equation methods tend to be less computationally expensive and therefore enjoyed much more attention, see for instance Margrave and Ferguson (1999); Margrave et al. (2006); Ma and Margrave (2008) and Al-Saleh et al. (2009).

Wavefield extrapolation methods, which image shot gathers independently, consist of three steps. First the incident wavefield is extrapolated either in depth for one way wave equation migration or in time for two way wave equation migration. The data are backpropaged from the surface into the model. Then an imaging condition is applied, where the forward propagated source wavefield and the back propagated data wavefield are put together to produce an image. Imaging conditions include the deconvolution imaging condition of Claerbout (1971) or the cross-correlation imaging condition which is often used specially for reverse time migration (Baysal et al., 1983; McMechan, 1983; Whitmore, 1983).

The Kirchhoff or integral methods are based on the concept of smearing the data over isochrones or summation over diffractions (Hagedoorn, 1954; Schneider, 1978; Miller et al., 1987). Isochrones are surfaces of constant cumulative time from the source to the image point to the receiver. The isochrone is an ellipsoid (ellipse in 2D) for homogeneous media (Miller et al., 1987).

Note that for historical reasons wavefield extrapolation methods are called wave equation methods. Kirchhoff methods are based on the wave equation but originally based on heuristic methods that were shown later to be based on the wave equation (Schneider, 1978; Bevc and Biondi, 2005).

1.5 Waveform inversion

Waveform inversion can be used to infer the Earth's properties including the compressional velocities. It is more general than seismic migration. In fact, migration as Lailly

(1983) and Tarantola (1984a) have shown, is roughly the first iteration of waveform inversion if all the frequencies are migrated. This is indeed the case because the first iteration of waveform inversion is roughly the migration of data residuals and the data residuals, if left without preconditioning, would be dominated by reflections for smooth background models (Shragge, 2007).

Waveform inversion can provide us with two main categories of useful information; an image of the subsurface and also a quantitative measure of the physical properties including velocities. Therefore, waveform inversion plays the role of migration and adds more to it the quantification of the physical properties. Mora (1989) summarizes this fact by the title of his paper, inversion = migration + tomography.

1.5.1 Complex near surface

The Earth's near surface can distort those images we make of the deeper part of the Earth. The near surface causes time shifts to the recored seismic data and therefore obscures the deeper part of the crust that is of interest to the upstream hydrocarbon industry (Sheriff and Geldart, 1995).

In order to correct for those distortions (time shifts), we often use some methods that attempt to infer the near surface velocities. Those methods, e.g. refraction methods, traveltime tomography etc. are robust and highly reliable, as it is well known (e.g. Yilmaz, 2001; Cox, 1999; Sheriff and Geldart, 1995). However they fall short when applied to areas with highly complex near surface geology. This is mainly due to the fact that they are based on an attribute of the wave field, which is traveltime. Waveform inversion, on the other hand, takes advantage the wavefield and attempts to directly invert the waveforms rather than an attribute of the waveforms. Keeping this in mind, it is not surprising then that waveform inversion is capable of creating such highly resolved models (Pratt, 1999).

The problems that arise from the near surface obviously depend on the complexity of the near surface. Areas with relatively simple near surface may not need waveform inversion since the existing methods do a good job. Areas with complex near surface on the other hand are the ones that will benefit the most of waveform inversion since the conventional methods are challenged. Therefore, we focus our attention particularly on complex near surface environments.

1.5.2 Difficulty of waveform inversion

Waveform inversion shares the same mentioned difficulties of inverse problems in general and has its own particular challenges. It is a highly non-linear problem and requires special data that may not be available such as large offsets and low frequencies (Sirgue and Pratt, 2004a). It is computationally expensive and therefore requires a very efficient modeling scheme. Frequency-domain waveform inversion in particular requires numerical forward modeling in the frequency domain, which is not a main stream modeling method, and therefore requires special expertise for implementation. This is in contrast to the often readily available time-domain modeling methods in any organization. Implementing an efficient modeling algorithm in particular is more difficult. Waveform inversion also requires an accurate initial model, relative to the available data frequencies. Such models are often difficult to obtain in practice (Schuster, 1998).

Waveform inversion attempts to invert for seismic velocities. Inverting for velocities is a very challenging problem in particular. As Symes (2009) mentions, in the last few decades, very little progress has been made with respect to velocity inversion as opposed to migration, for instance. In fact, some methods have been proposed and fall under the category of migration velocity analysis (e.g. Al-Yahya, 1989; Biondi and Sava, 1999; Mosher et al., 2001) including the differential semblance method (e.g. Symes and Carazzone, 1991; Plessix, 2000; Brandsberg-Dahl et al., 2003; Shen and Symes, 2008)

and stereotomography (e.g. Billette and Lambaré, 1998; Billette et al., 2003; Lambaré, 2008). But they all come with their own limitations and therefore are still under active research.

1.6 Thesis objectives

The objectives of this thesis are both theoretical and practical. We attempt to put waveform inversion in context with other methods of imaging and inversion. Much of this work has been published already but very few attempts have been made to put most of it together in a coherent fashion. The other objective of the thesis is to study the behavior of waveform inversion for areas with very challenging near surface problems. We study how waveform inversion behaves for very heterogeneous near surface models. We apply it to an area of very complex near surface, namely Saudi Arabia. Finally, we study what waveform inversion can offer to solve the challenging problem of first arrival shingling.

1.7 Thesis overview and contributions

The first chapter, the current one, is an overview of inversion, migration and specifically our main interest, waveform inversion. In the second chapter, we discuss the theory of waveform inversion and put it into context with other existing methods since the theory is scattered in the literature for almost three decades. We derive, discuss and interpret some properties of waveform inversion. We do that while trying to explicitly show the embedded assumptions in waveform inversion and how it relates to migration in general, Kirchhoff migration, traveltime tomography and true amplitude migration/inversion.

We address in the third chapter the heterogeneity of the near surface and its effects on waveform inversion. Complex near surfaces are obviously very heterogeneous;

unconsolidated sediments, caves, gravels boulders mixed with soil/sand and so on all represent extreme velocity variations. We study how waveform inversion can address such challenges and we show that it can handle such extreme variations.

We focus our attention in the fourth chapter on a realistic environment with very complex near surface, namely that in Saudi Arabia, which is very well known for its challenging near surface. We show how waveform inversion can address the challenges associated with such areas. We evaluate the data requirement and weigh the potential benefits. We also point out the potential pitfalls of waveform inversion for such environments.

In the fifth chapter, we address a specific long lasting challenge, namely first arrival shingling. First arrival shingling is a particular problem that is associated with some challenges in the near surface. As we will show, research into this problem has been conducted since the 1950s, yet no generic solution exists, as far as we are aware. We attempt to investigate what can waveform inversion can do to address this issue and we show that waveform inversion can indeed resolve such a challenge.

1.8 Tools used

The results and the figures in the thesis have been created using different tools. The schematic figures have been created using Microsoft Office member, Powerpoint. I have used the ubiquitous scientific software package, Matlab from MathWorks, including the CREWES MatLab toolbox authored by Dr. Gary Margrave and coworkers, at the University of Calgary CREWES project. The open source software package, Seismic Un*x, from the Center for Wave Phenomena at Colorado School of Mines, is also used. I have also used frequency-domain finite-difference modeling and inversion packages by Dr. Gerhard Pratt, currently at the University of Western Ontario, and his coworkers.

Typesetting the thesis was done using LaTeX.

This page is intentionally left blank.

Chapter 2

Theory of waveform inversion

The goal of this chapter is to give a coherent theoretical treatment for this subject, waveform inversion. We attempt to glue together various methods from various sources in the literature. We discuss the details of frequency-domain finite-difference modeling. We derive the fundamental equation in waveform inversion. Namely we derive the gradient of the misfit function using the Born approximation and using matrix formulation, as well. We compare and contrast waveform inversion to the mainstream seismic imaging method, seismic migration. We compare it to Kirchhoff migration, true amplitude inversion/migration, as well as, traveltime tomography. We, also, detail the different steps involved in regularizing and alleviating the non-linearity of the problem in order to obtain a reasonable model.

2.1 Introduction

Inversion is a process of inferring the model parameters from the given data. Before seeking the solution to the inverse problem, the forward problem needs to be formulated first (Aster et al., 2005; Tarantola, 2005). Therefore, we discuss the forward problem, first, which is, in our case, frequency-domain finite-difference modeling. We also discuss the associated techniques that make solving such a problem not only efficient but also feasible for such realistic models. The inverse problem is then investigated with the aim of gaining some insight and understanding of its mechanics.

Inverting the waveforms is a highly non-linear process Sirgue (2003). Due to multiple scattering, perturbations in the data do not map linearly to perturbations in the model.

However, we can always linearize. The linearization of this problem is what's commonly referred to as the first-order Born approximation in physics. As we will see, the validity of the Born approximation stems from the assumption that the model perturbation is too weak to generate any significant multiples. In the context of seismic imaging and inversion, this entails that we know very closely the background model and then our task will be to estimate the model perturbation. Although this is a strict limitation to migration and inversion, there exists some workarounds.

Waveform inversion is based on solving a highly non-linear problem by linearized iterative inversion. The formula is linearized and then we approach the solution through iteration. In other words, non-linear inversion is achieved by solving a linear problem and then basing the next iteration of the previous result and so on.

2.2 Convention

Before diving into the theory, we will start by stating the conventions we will be using. Since Bleistein et al. (2001) provide one of the most rigorous yet pragmatic presentations of inverse scattering theory, we will try to follow their convention with slight differences where necessary. To start with, we will use their convention of the Fourier transform. The temporal Fourier forward and inverse transforms we will use are

$$f(\omega) = \int_{-\infty}^{\infty} F(t) e^{i\omega t} dt \qquad Forward \tag{2.1}$$

and

$$F(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(\omega) e^{-i\omega t} d\omega \qquad Inverse, \tag{2.2}$$

where t is the time, and $\omega$is the frequency. The spatial forward and inverse Fourier transforms are

$$f(k) = \int_{-\infty}^{\infty} F(x) e^{-ikx} dx \qquad Forward \tag{2.3}$$

and

$$F(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(k)e^{ikx}dk \qquad Inverse, \tag{2.4}$$

where k is the wavenumber and x is a spatial coordinate.

2.3 Forward modeling

Forward modeling is an integral part of waveform inversion. As we will see, the inverse problem uses the adjoint of forward modeling, which is migration. Therefore, an efficient modeling method is necessary in order to make the problem tractable.

In principle, any method can be used to carry out the forward modeling (Mora, 1987). Methods such as the reflectivity method, Kirchhoff modeling, WKBJ method, Maslov method, Gaussian beam, (Chapman, 2004; Aki and Richards, 2002) and numerical methods such as finite-difference, finite-element can all be used. However, any limitations of the forward problem will certainly impose limitations on the inversion (Mora, 1987).

Frequency-domain finite-difference modeling is a very efficient method of modeling specifically for waveform inversion (Štekl and Pratt, 1998). As we will discuss, solving the inverse problem only requires modeling and inverting a number of frequencies.

2.3.1 Modeling equations

We start with the scalar, acoustic, variable density wave equation in the frequency-domain

$$\nabla\left(\frac{1}{\rho(\mathbf{x})}\nabla u(\mathbf{x},\mathbf{x}_s,\omega)\right) + \frac{\omega^2}{v^2(\mathbf{x})\rho(\mathbf{x})}u(\mathbf{x},\mathbf{x}_s,\omega) = -f(\omega)\delta(\mathbf{x}-\mathbf{x}_s), \tag{2.5}$$

where $u(\mathbf{x},\mathbf{x}_s,\omega)$ is the wavefield due to a source located at $\mathbf{x}_s$, $f(\omega)$ is the source signature, $\rho(\mathbf{x})$ and $v(\mathbf{x})$ are the density and the velocity of the medium, respectively (Bleistein et al., 2001, p.94). The equation can also be written as

$$\left[\nabla\left(\frac{1}{\rho(\mathbf{x})}\nabla\right) + \frac{\omega^2}{v^2(\mathbf{x})\rho(\mathbf{x})}\right] u(\mathbf{x},\mathbf{x_s},\omega) = -f(\omega)\delta(\mathbf{x}-\mathbf{x}_s). \tag{2.6}$$

By letting

$$\mathcal{A} = \left[\nabla\left(\frac{1}{\rho(\mathbf{x})}\nabla\right) + \frac{\omega^2}{v^2(\mathbf{x})\rho(\mathbf{x})}\right], \tag{2.7}$$

we can write the equation 2.6 as

$$\mathcal{A}u(\mathbf{x}, \mathbf{x_s}, \omega) = -f(\omega)\delta(\mathbf{x} - \mathbf{x}_s), \tag{2.8}$$

where $\mathcal{A}$ is the forward modeling operator. For constant density media, equation 2.6 reduces to

$$\left[\nabla^2 + \frac{\omega^2}{v^2(\mathbf{x})}\right] u(\mathbf{x}, \mathbf{x}_s, \omega) = -f(\omega)\delta(\mathbf{x} - \mathbf{x}_s), \tag{2.9}$$

which is what is commonly referred to as the Helmholtz equation.

An advantage of modeling in the frequency domain is the simplicity of introducing attenuation; we simply make the velocities complex (Štekl, 1997). The equation implemented is

$$v(\mathbf{x}) = v_o(\mathbf{x})\left[1 - i\frac{2}{Q}\right], \tag{2.10}$$

where Q is the quality factor (Brenders and Pratt, 2007b; Pratt, 2008).

2.3.2 Frequency-domain finite-difference (FDFD) modeling

When descritizing the wave equation 2.8, the equation can be written in matrix form as

$$\mathbf{A}(\mathbf{x}, \mathbf{x}_s, \omega)\mathbf{u}(\mathbf{x}, \mathbf{x}_s, \omega) = \mathbf{f}(\omega), \tag{2.11}$$

where $\mathbf{A}$ is the forward modeling operator matrix, $\mathbf{u}$ is the wavefield vector and $\mathbf{f}$ is the source vector.

Equation 2.11 is solved using a direct method, LU decomposition (Pratt, 2008). Although an iterative method might seem more efficient, currently, there exists no generic method to solve this problem iteratively (Erlangga, 2008). It is an active area of research

in applied mathematics. However, some iterative methods have been proposed (e.g. Erlangga et al., 2004; Plessix, 2006) and some have been applied, only recently, to real data (Plessix, 2009).

The forward operator matrix is a sparse matrix (Pratt et al., 1998). LU decomposition is used to factorize the matrix, but produces matrices with more elements than in the original matrix. This requires a lot of computer memory. To alleviate this issue Štekl (1997); Štekl and Pratt (1998) applied nested dissection. Nested dissection permutes/re-arranges the elements of the matrix in such a way that the resulting matrices have minimum number of elements and therefore require less computer memory. The implementation of FDFD is detailed in Pratt (1990); Jo et al. (1996); Štekl and Pratt (1998).

The efficiency of FDFD for waveform inversion stems from a number of points:

1. Only a number of frequencies need to be modeled and inverted for waveform inversion.

2. If direct matrix inversion methods are used, multiple shots can be modeled efficiently since the forward operator matrix needs to be inverted/factored only once.

3. Inverting the low frequencies (e.g. < 10 Hz) requires sparser grid, although this point applies to time-domain-finite difference, as well.

2.3.3 Boundary conditions

Absorbing boundary conditions are implemented on the boundaries of the model. The 45 degree one-way wave equation is used (Clayton and Engquist, 1977). At the corners of the model, an additional sponge-like boundary condition is used with a finite value of Q. But in order for the corner boundary condition to work effectively, the receiver needs to be at least 10 grid points away from the corner (Brenders and Pratt, 2007b).

2.3.4 Time domain synthetics

Since the synthetic data are generated as discrete frequencies in the frequency domain, transforming them into the time domain results in aliased seismograms. This issue is addressed by Mallick and Frazer (1987) and also explained by Brenders and Pratt (2007b). Here, we discuss the details of the method.

Let $u(\omega)$ be the seismic trace in the frequency domain. When transforming $u(\omega)$ to the time domain, we don't get $U(t)$, but rather, we get

$$U_w(t) = \sum_{n=-\infty}^{\infty} U(t + nT_{max}), \tag{2.12}$$

which is a version of the trace with aliasing or wraparound. To attenuate aliasing, two steps are involved:

1. Make the frequency imaginary by adding a small $\frac{1}{\tau}$ to the imaginary part. That is let $u_c(\omega') = u(\omega + i\frac{1}{\tau})$.

2. Transform u_c to time domain to get U_c. Multiply U_c by $e^{\frac{t}{\tau}}$ in order to obtain U_c'.

To see why this works, let's start with the first step. Making the frequency complex by adding an imaginary term $\frac{1}{\tau}$ to any signal results in

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega e^{-i\omega t} u(\omega + i\frac{1}{\tau}) = U(t)e^{-t/\tau}, \tag{2.13}$$

when transformed to the time domain. Therefore, we get the signal in the time domain but it is damped by the factor $e^{-t/\tau}$. In discrete form equation 2.13 becomes

$$\frac{\Delta\omega}{2\pi}\sum_{n=-\infty}^{\infty} e^{-int\Delta\omega} u(n\Delta\omega + i\frac{1}{\tau}) = \sum_{n=-\infty}^{\infty} U(t + nT_{max})e^{-(t+nT_{max})/\tau} = U_c. \tag{2.14}$$

To undo the damping factor we multiply the U_c by $e^{t/\tau}$. That is

$$U_c' = U_c e^{t/\tau} = \sum_{n=-\infty}^{\infty} U(t+nT_{max}) e^{-(t+nT_{max})/\tau} e^{t/\tau} = \sum_{n=-\infty}^{\infty} U(t+nT_{max}) e^{-nT_{max}/\tau}, \quad (2.15)$$

which is

$$U_c'(t) = U(t) + \sum_{n\neq 0} U(t+nT_{max}) e^{-n\frac{T_{max}}{\tau}}. \quad (2.16)$$

Therefore, we obtained the original signal $U(t)$ plus some wraparound. The wraparound can be effectively attenuated, then, by choosing a proper τ.

2.4 Waveform inversion

Waveform inversion is an iterative linearized inversion Tarantola (1984a). The Born approximation is embedded into the method. We will derive the formulas for waveform inversion to illustrate these points. We will use two methods. The first is based on that of Tarantola (1984a) and Mora (1987) but in the frequency domain. The second is based on Pratt et al. (1996) and we will compare the two methods. We first state the tools we will need for such derivation.

2.4.1 Born approximation

We state in this section the integral solution to the wave equation, the so called the Lippmann-Schwinger integral and its linearization, which is referred to as the Born approximation. We follow in this section the treatment given in Bleistein et al. (2001).

We state here, again, the Helmholtz equation,

$$\left[\nabla^2 + \frac{\omega^2}{v^2(\mathbf{x})}\right] u(\mathbf{x}, \mathbf{x}_s, \omega) = -f(\omega)\delta(\mathbf{x} - \mathbf{x}_s) \quad (2.17)$$

The Green's function $g(\mathbf{x}, \mathbf{x}_s, \omega)$ or the impulse response is the solution to the wave equation when the source term is the delta function

$$\left[\nabla^2 + \frac{\omega^2}{v^2(\mathbf{x})}\right] g(\mathbf{x}, \mathbf{x}_s, \omega) = -\delta(\mathbf{x} - \mathbf{x}_s) \tag{2.18}$$

What is unique about the Green's function is that once it is known, the wavefield $u(\mathbf{x}, \mathbf{x}_s, \omega)$ due to any arbitrary source wavelet can be obtained by

$$u(\mathbf{x}, \mathbf{x}_s, \omega) = f(\omega) g(\mathbf{x}, \mathbf{x}_s, \omega), \tag{2.19}$$

Let

$$\frac{1}{v^2(\mathbf{x})} = m(\mathbf{x}) = m_0(\mathbf{x}) + \delta m(\mathbf{x}), \tag{2.20}$$

where $m_0(\mathbf{x})$ is the background model and $\delta m(\mathbf{x})$ is a model perturbation. Also, let

$$u_{obs}(\mathbf{x}, \mathbf{x_s}, \omega) = u(\mathbf{x}, \mathbf{x}_s, \omega) + \delta u(\mathbf{x}, \mathbf{x}_s, \omega), \tag{2.21}$$

where $u_{obs}(\mathbf{x}, \mathbf{x_s}, \omega)$ is the observed wavefield, $u(\mathbf{x}, \mathbf{x}_s, \omega)$ is the wave field in the background model m_0 and $\delta u(\mathbf{x}, \mathbf{x}_s, \omega)$ is the wavefield due to the perturbation $\delta m(\mathbf{x})$ which can be thought of as the scattered wavefield. $u_{obs}(\mathbf{x}, \mathbf{x_s}, \omega)$ is the total wavefield, In other words, it is the sum of the wavefield in the background model and the scattered wavefield.

An integral solution to the wave equation for the scattered wavefield is

$$\delta u(\mathbf{x_g}, \mathbf{x_s}, \omega) = \omega^2 \int_V \delta m(\mathbf{x}) \left[u(\mathbf{x}, \mathbf{x_s}, \omega) + \delta u(\mathbf{x}, \mathbf{x_s}, \omega)\right] g(\mathbf{x_g}, \mathbf{x}, \omega) d^3x, \tag{2.22}$$

which is often referred to as the Lippman-Schwinger integral formula, see (e.g. Bleistein et al., 2001) for the derivation. The integral contains the total wavefield which includes the scattered wavefield. This makes it nonlinear. If we assume that the model perturbation, δm, is sufficiently weak, then we can drop the scattered wavefield term, $\delta u(\mathbf{x}, \mathbf{x_s}, \omega)$, and therefore linearize equation 2.22. This results in

$$\delta u(\mathbf{x_g}, \mathbf{x_s}, \omega) = \omega^2 \int_V \delta m(\mathbf{x}) u(\mathbf{x}, \mathbf{x_s}, \omega) g(\mathbf{x_g}, \mathbf{x}, \omega) d^3x. \tag{2.23}$$

This is what is referred to as the Born approximation. If we let

$$u(\mathbf{x}, \mathbf{x}_s, \omega) = f(\omega) g(\mathbf{x}, \mathbf{x}_s, \omega), \tag{2.24}$$

equation 2.23 can also be written as

$$\delta u(\mathbf{x}_g, \mathbf{x}_s, \omega) = \omega^2 f(\omega) \int_V \delta m(\mathbf{x}) g(\mathbf{x}, \mathbf{x}_s, \omega) g(\mathbf{x}_g, \mathbf{x}, \omega) d^3x. \tag{2.25}$$

2.4.2 Derivation using non-matrix formulation

We derive here the formula for waveform inversion using non-matrix methods. This derivation follows that of Tarantola (1984a) and Mora (1987) but in the frequency domain. It is similar to that of Schuster (1998). We begin with defining the objective function. The objective or misfit function is the L2 norm of the difference between the observed waveforms and the modeled ones, as follow:

$$E = \frac{1}{2} \int d\omega \sum_s \sum_g |\delta u(\mathbf{x}_g, \mathbf{x}_s, \omega)|^2, \tag{2.26}$$

where

$$\delta u(\mathbf{x}_g, \mathbf{x}_s, \omega) = u_{obs}(\mathbf{x}_g, \mathbf{x}_s, \omega) - u(\mathbf{x}_g, \mathbf{x}_s, \omega), \tag{2.27}$$

which is the data residuals. We seek an iterative scheme to update the model by letting

$$m(\mathbf{x})^{(n+1)} = m(\mathbf{x})^{(n)} + \mu^{(n)} \gamma(\mathbf{x})^{(n)} \tag{2.28}$$

where

$$\gamma(\mathbf{x}) = \nabla_m E = \frac{\partial E}{\partial m(\mathbf{x})}, \tag{2.29}$$

which is the gradient of the misfit function E and μ is the step length. Evaluating the gradient, we get

$$\gamma(\mathbf{x}) = \int d\omega \sum_s \sum_g Re[\frac{\partial u(\mathbf{x}_g, \mathbf{x}_s, \omega)}{\partial m(\mathbf{x})} \delta u^*(\mathbf{x}_g, \mathbf{x}_s, \omega)]. \tag{2.30}$$

where $\delta u^*(\mathbf{x}_g, \mathbf{x}_s, \omega)$ is the complex conjugate of $\delta u(\mathbf{x}_g, \mathbf{x}_s, \omega)$.

In order to find the Fréchet derivative $\frac{\partial u(\mathbf{x}_g, \mathbf{x}_s, \omega)}{\partial m(\mathbf{x})}$, we seek an operator $\mathbf{D}$ such that

$$\delta \mathbf{d} = \mathbf{D} \delta \mathbf{m}, \tag{2.31}$$

which maps perturbations in the data $\delta \mathbf{d}$ to perturbations in the model $\delta \mathbf{m}$. In continuous form, this becomes

$$\delta d = \int_V d^3x \, \frac{\partial u(\mathbf{x}_g, \mathbf{x}_s, \omega)}{\partial m(\mathbf{x})} \delta m. \tag{2.32}$$

The integral formula, which maps perturbations in the model space and to those in the data space, is what we have seen previously, namely the Born integral formula 2.25. We state it here again

$$\delta u(\mathbf{x}_g, \mathbf{x}_s, \omega) = \int_V d^3x \, \omega^2 f(\omega) g(\mathbf{x}_g, \mathbf{x}, \omega) g(\mathbf{x}, \mathbf{x}_s, \omega) \delta m(\mathbf{x}) \tag{2.33}$$

Therefore,

$$\frac{\partial u(\mathbf{x}_g, \mathbf{x}_s, \omega)}{\partial m(\mathbf{x})} = \omega^2 f(\omega) g(\mathbf{x}_g, \mathbf{x}, \omega) g(\mathbf{x}, \mathbf{x}_s, \omega). \tag{2.34}$$

Substituting equation 2.34 into equation 2.30, the gradient of the misfit function then becomes:

$$\gamma(\mathbf{x}) = \int d\omega \sum_s \sum_g Re[\omega^2 f(\omega) \delta u^*(\mathbf{x}_g, \mathbf{x}_s, \omega) g(\mathbf{x}_g, \mathbf{x}, \omega) g(\mathbf{x}, \mathbf{x}_s, \omega)]. \tag{2.35}$$

We can interpret this formula as

$$\gamma(\mathbf{x}) = \int d\omega \sum_s \sum_g Re[\omega^2 \overbrace{f(\omega)g(\mathbf{x},\mathbf{x}_s,\omega)}^{\text{forward propagated wavefield}} \underbrace{\delta u^*(\mathbf{x}_g,\mathbf{x}_s,\omega)g(\mathbf{x}_g,\mathbf{x},\omega)}_{\text{backpropagated wavefield}}]. \quad (2.36)$$

In other words, it is the time domain zero-large cross-correlation between the forward propagated wavefield and the back propagated wavefield (Lailly, 1983; Tarantola, 1984a) just as in done in reverse time migration (e.g. Pratt et al., 1998).

Although the notation in equation 2.35 implies that we model each receiver Green's function $g(\mathbf{x}_g,\mathbf{x},\omega)$ independently for every receiver, in reality we don't need to do that. We can model all of them simultaneously. This point is often not discussed in sufficient detail. Here, we will see why it is true.

This point is valid because of the summation over the receivers for the objective function and hence the gradient of the misfit function. The summation over the receivers makes this point valid because of the linearity of the wave equation. We will illustrate this point for two receivers at locations $\mathbf{x}_{g_1}$ and $\mathbf{x}_{g_2}$. If we fix the shots and the frequencies for simplicity and drop the real part notation, the gradient of the misfit in equation 2.35 for only two receivers then becomes

$$\gamma(\mathbf{x}) = \omega^2 f(\omega)\delta u^*(\mathbf{x}_{g_1},\omega)g(\mathbf{x},\mathbf{x}_s,\omega)g(\mathbf{x}_{g_1},\mathbf{x},\omega)+\omega^2 f(\omega)\delta u^*(\mathbf{x}_{g_2},\mathbf{x}_s,\omega)g(\mathbf{x},\mathbf{x}_s,\omega)g(\mathbf{x}_{g_2},\mathbf{x},\omega). \quad (2.37)$$

For simplicity, we dropped the dependence of δu^*on $\mathbf{x_s}$. By letting

$$u_b(\mathbf{x}_{g_1},\mathbf{x},\omega) = \delta u^*(\mathbf{x}_{g_1},\omega)g(\mathbf{x}_{g_1},\mathbf{x},\omega) \quad (2.38)$$

and

$$u_b(\mathbf{x}_{g_2},\mathbf{x},\omega) = \delta u^*(\mathbf{x}_{g_2},\omega)g(\mathbf{x}_{g_2},\mathbf{x},\omega), \quad (2.39)$$

we can write the gradient as

$$\gamma(\mathbf{x}) = \omega^2 f(\omega) g(\mathbf{x}, \mathbf{x}_s, \omega)[u_b(\mathbf{x}_{g_1}, \mathbf{x}, \omega) + u_b(\mathbf{x}_{g_2}, \mathbf{x}, \omega)]. \tag{2.40}$$

So, that letting

$$\tilde{u_b}(\mathbf{x}_{g_1}, \mathbf{x}_{g_2}, \mathbf{x}, \omega) = u_b(\mathbf{x}_{g_1}, \mathbf{x}, \omega) + u_b(\mathbf{x}_{g_2}, \mathbf{x}, \omega), \tag{2.41}$$

then we can simplify equation 2.40 further to

$$\gamma(\mathbf{x}) = \omega^2 f(\omega) g(\mathbf{x}, \mathbf{x}_s, \omega) \tilde{u_b}(\mathbf{x}_{g_1}, \mathbf{x}_{g_2}, \mathbf{x}, \omega). \tag{2.42}$$

We will seek a formula for $\tilde{u_b}(\mathbf{x},\mathbf{x}_{g_1}, \mathbf{x}_{g_2}, \omega)$ by taking advantage of the linearity of the wave equation. Using the wave equation

$$\mathcal{L}u(\mathbf{x}_{g_1}, \mathbf{x}, \omega) = -f(\omega)\delta(\mathbf{x} - \mathbf{x}_{g_1}), \tag{2.43}$$

by the definition of equation 2.39 the source signature $f(\omega)$, then becomes

$$f(\omega) = \delta u^*(\mathbf{x}_{g_1}, \mathbf{x}_s, \omega). \tag{2.44}$$

Then, we can write equation 2.43 as

$$\mathcal{L}u(\mathbf{x}_{g_1}, \mathbf{x}, \omega) = -\delta u^*(\mathbf{x}_{g_1}, \omega)\delta(\mathbf{x} - \mathbf{x}_{g_1}). \tag{2.45}$$

In a similar manner,

$$\mathcal{L}u(\mathbf{x}_{g_2}, \mathbf{x}, \omega) = -\delta u^*(\mathbf{x}_{g_2}, \omega)\delta(\mathbf{x} - \mathbf{x}_{g_2}). \tag{2.46}$$

Adding equations 2.45 and 2.46 has the form

$$\mathcal{L}[u(\mathbf{x}_{g_1}, \mathbf{x}, \omega) + u(\mathbf{x}_{g_2}, \mathbf{x}, \omega)] = -\delta u^*(\mathbf{x}_{g_1}, \omega)\delta(\mathbf{x} - \mathbf{x}_{g_1}) - \delta u^*(\mathbf{x}_{g_2}, \omega)\delta(\mathbf{x} - \mathbf{x}_{g_2}). \tag{2.47}$$

Using the definition of $\tilde{u_b}(\mathbf{x},\mathbf{x}_{g_1},\mathbf{x}_{g_2},\omega)$ in equation 2.41, the above equation can be written as

$$\mathcal{L}\tilde{u_b}(\mathbf{x},\mathbf{x}_{g_1},\mathbf{x}_{g_2},\omega) = -\delta u^*(\mathbf{x}_{g_1},\omega)\delta(\mathbf{x}-\mathbf{x}_{g_1}) - \delta u^*(\mathbf{x}_{g_2},\omega)\delta(\mathbf{x}-\mathbf{x}_{g_2}), \tag{2.48}$$

Equation 2.48 shows that we can model the wavefields at all receivers simultaneously by making the adjoint of the data residuals as sources. In other words, there is no need to model the wavefields at every receiver separately. This saves us a big number of unnecessary forward modeling iterations. It saves for a single shot as many forward modeling iterations as the number of data points plus an additional modeling iteration for the source.

2.4.3 Derivation using matrix formulation

We follow here the derivation given in Pratt et al. (1998). We state again the modeling equation for one single shot in the frequency-domain

$$\mathbf{Au} = \mathbf{f}, \tag{2.49}$$

where $\mathbf{A}$ is the modeling operator matrix, $\mathbf{u}$ is the vector containing the wavefield at every grid point and $\mathbf{f}$ is source vector. Our objective again is to minimize some function that involves the difference between the observed data, $\mathbf{d}$, and the synthetic data, $\mathbf{u}$. This difference, which is given by

$$\delta\mathbf{d} = \mathbf{d} - \mathbf{u}. \tag{2.50}$$

To be more precise we should let $\mathbf{u}$ be $\mathbf{Ru}$, where $\mathbf{R}$ is a matrix that restricts the synthetic wavefield vector to the observed data points only and we should let$\mathbf{d}$ be $\mathbf{Pd}$, where $\mathbf{P}$ pads the data vector with zeros where observation points do not exist (Virieux

and Operto, 2009). However for simplicity, we dropped them. The objective function or misfit function is the L2 norm of the data residuals

$$E = \frac{1}{2} \sum_{\omega} \sum_{s} \delta \mathbf{d^t} \delta \mathbf{d}^*. \tag{2.51}$$

An iterative method like the steepest decent would look like

$$\mathbf{m^{(n+1)}} = \mathbf{m^{(n)}} - \mu^{\mathbf{n}} \gamma^{\mathbf{n}}, \tag{2.52}$$

where $\gamma = \nabla_{\mathbf{m}} \mathbf{E} = \frac{\partial \mathbf{E}}{\partial \mathbf{m}}$. So, we start by finding the gradient of the misfit function,$\nabla_m E$, which is

$$\nabla_m E = \frac{\partial E}{\partial \mathbf{m}} = Re\{\mathbf{J}^t \delta d^*\}, \tag{2.53}$$

where $\mathbf{J}$ is the Fréchet derivative matrix, which can also be written as

$$J_{ij} = \frac{\partial u_i}{\partial m_j}, \tag{2.54}$$

where $i = 1, 2, .., n$ where n is the number of receivers and $j = 1, 2, 3, .., m$, where m is the number of model parameters. The Fréchet derivative matrix is a huge matrix that requires a big number of forward modeling iterations. This number is roughly equal to the number of model parameters for every shot. In order to better visualize the matrices we have them here written explicitly, just as shown in Pratt et al. (1998)

$$\begin{bmatrix} \frac{\partial E}{\partial m_1} \\ \frac{\partial E}{\partial m_2} \\ \vdots \\ \frac{\partial E}{\partial m_l} \end{bmatrix} = Re \left\{ \begin{bmatrix} \frac{\partial u_1}{\partial m_1} & \frac{\partial u_2}{\partial m_1} & \cdots & \frac{\partial u_l}{\partial m_1} \\ \frac{\partial u_1}{\partial m_2} & \frac{\partial u_2}{\partial m_2} & \cdots & \frac{\partial u_l}{\partial m_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial u_1}{\partial m_l} & \frac{\partial u_2}{\partial m_l} & \cdots & \frac{\partial u_l}{\partial m_l} \end{bmatrix} \begin{bmatrix} \delta d_1^* \\ \delta d_2^* \\ \vdots \\ \delta d_l^* \end{bmatrix} \right\}. \tag{2.55}$$

We can lump together the columns of equation 2.55 and write it as

$$\begin{bmatrix} \frac{\partial E}{\partial m_1} \\ \frac{\partial E}{\partial m_2} \\ \vdots \\ \frac{\partial E}{\partial m_l} \end{bmatrix} = Re \left\{ \begin{bmatrix} \frac{\partial \mathbf{u}^t}{\partial m_1} \\ \frac{\partial \mathbf{u}^t}{\partial m_2} \\ \vdots \\ \frac{\partial \mathbf{u}^t}{\partial m_l} \end{bmatrix} \begin{bmatrix} \delta d_1^* \\ \delta d_2^* \\ \vdots \\ \delta d_l^* \end{bmatrix} \right\}.$$

Therefore we can write the formula for the Fréchet derivatives in a more compact form as

$$\mathbf{J_i} = \frac{\partial \mathbf{u}}{\partial m_i}. \tag{2.56}$$

However, in order to find a more efficient solution, we first start by taking the derivative of the forward modeling equation 2.49

$$\mathbf{A}\frac{\partial \mathbf{u}}{\partial m_i} = -\frac{\partial \mathbf{A}}{\partial m_i}\mathbf{u},$$

which can be re-written as

$$\frac{\partial \mathbf{u}}{\partial m_i} = -\mathbf{A}^{-1}\frac{\partial \mathbf{A}}{\partial m_i}\mathbf{u}. \tag{2.57}$$

Substituting equation 2.56 into 2.57, we obtain

$$\mathbf{J_i} = \mathbf{A}^{-1}\frac{\partial \mathbf{A}}{\partial m_i}\mathbf{u}. \tag{2.58}$$

In order to obtain an expression for the gradient of the misfit function, we substitute 2.58 into 2.53 to obtain

$$\frac{\partial E}{\partial m_i} = Re\{\mathbf{J}_i^t \delta d^*\} = Re\{- \underbrace{\mathbf{u}^t \frac{\partial \mathbf{A}^t}{\partial m_i}}_{\text{Forward wavfield}} \underbrace{\mathbf{A}^{-1}\delta d^*}_{\text{Backprop. wavefiled}} \}. \tag{2.59}$$

Notice that the matrix $\frac{\partial \mathbf{A}^t}{\partial m_i}$ has obviously zeros everywhere except at the elements that depend on m_i. In addition the matrix $\mathbf{A}$ itself is a sparse matrix as mentioned previously.

Therefore, equation 2.59 for all model elements is essentially equivalent to equation 2.36 but in matrix form (Pratt et al., 1998).

2.4.4 Analytic example

In order to gain further insight into the machinery of waveform inversion, we will look at an analytic example. We will use the far field approximation and therefore approximate the Green's functions with plane waves and see what waveform inversion does to every plane wave component. We follow here the derivation by Sirgue and Pratt (2004a).

So let the source Green's function be

$$g(\mathbf{x},\omega) \approx e^{ik_o \hat{\mathbf{r}}_s \cdot \mathbf{x}}, \tag{2.60}$$

and the receiver Green's function

$$g(\mathbf{x},\omega) \approx e^{ik_o \hat{\mathbf{r}}_g \cdot \mathbf{x}}, \tag{2.61}$$

where $\hat{\mathbf{r}}_s$ and $\hat{\mathbf{r}}_g$ are the unit vectors that point from the source to the scatterer or image point and from the receiver to the image point, respectively and $k_0 = \frac{\omega}{v}$. Substituting those Green's functions into the formula for the gradient of the misfit function (equation 2.35), we get for a single temporal frequency

$$\gamma(\mathbf{x}) = \int d\mathbf{x}_s \int d\mathbf{x}_g \, Re[\omega^2 F(\omega) \delta u^*(\hat{\mathbf{r}}_g, \hat{\mathbf{r}}_s, \omega) e^{ik_o \hat{\mathbf{r}}_s \cdot \mathbf{x}} e^{ik_o \hat{\mathbf{r}}_g \cdot \mathbf{x}}], \tag{2.62}$$

re-arranging, we get

$$\gamma(\mathbf{x}) = \omega^2 F(\omega) \int d\mathbf{x}_s \int d\mathbf{x}_g \, Re[\delta u^*(\hat{\mathbf{r}}_g, \hat{\mathbf{r}}_s, \omega) e^{ik_o \mathbf{x} \cdot (\hat{\mathbf{r}}_s + \hat{\mathbf{r}}_g)}]. \tag{2.63}$$

We will seek now a formula for the data residuals $\delta u^(\hat{\mathbf{r}}_g, \hat{\mathbf{r}}_s, \omega)$ so that we could invoke the model perturbation into equation 2.63. In order to do that, we plug in the plane waves into the Born integral formula, equation 2.25, and we get

$$\delta u(\mathbf{x}_g, \mathbf{x}_s, \omega) = \omega^2 f(\omega) \int_V \delta m(\mathbf{x}) e^{ik_o \mathbf{x} \cdot (\hat{\mathbf{r}}_s + \hat{\mathbf{r}}_g)} d^3\mathbf{x}. \tag{2.64}$$

Notice that the integral in equation 2.64 is nothing but the Fourier transform of the model perturbation $\delta m(\mathbf{x})$. Therefore, we can write equation 2.64 as

$$\delta u(\mathbf{x}_g, \mathbf{x}_s, \omega) = \omega^2 f(\omega) \delta M(k_o(\hat{\mathbf{r}}_s + \hat{\mathbf{r}}_g)), \tag{2.65}$$

where $\delta M(k_o(\hat{\mathbf{r}}_s + \hat{\mathbf{r}}_g))$ is model perturbation in the wavenumber-domain.

Now that we sought an expression for the data residuals δu that invokes the perturbation, we will plug in this formula into the gradient of the misfit function to get

$$\gamma(\mathbf{x}) = \omega^2 f(\omega) \int d\mathbf{x}_s \int d\mathbf{x}_g \, Re[\left[\omega^2 f(\omega) \delta M(k_o(\hat{\mathbf{r}}_s + \hat{\mathbf{r}}_g)\right]^* e^{ik_o \mathbf{x} \cdot (\hat{\mathbf{r}}_s + \hat{\mathbf{r}}_g)}]. \tag{2.66}$$

Note that equation 2.66 is nothing but the inverse Fourier transform of the model perturbation and therefore, we can write equation it as

$$\gamma(\mathbf{x}) = \omega^4 f^2(\omega) \delta m(\mathbf{x}). \tag{2.67}$$

Therefore, the gradient of the misfit function was able to recover the model perturbation $\delta m(\mathbf{x})$ except that it is off by a complex scalar $f^2(\omega)$, which is the source strength, and multiplied by the frequency raised to the fourth power, ω^4. As Sirgue and Pratt (2004a) point out, the term ω^4 can be interpreted as the derivative operator, which is a roughening operator, raised to the fourth power. That is

$$\gamma(\mathbf{x}) = \underbrace{(-i\omega)(-i\omega)(-i\omega)(-i\omega)}_{\text{roughening operator}} f^2(\omega) \delta m(\mathbf{x}). \tag{2.68}$$

This shows the necessity of preconditioning the gradient of the misfit function since it contains some rough unnecessary features. We will discuss this point further in section 2.6.3.

2.4.5 Waveform inversion and Kirchhoff migration

We show in this section the relation between Kirchhoff migration and waveform inversion. This was first explained by Tarantola (1984b). We follow here the derivation of Tarantola (1984b) and Schuster (1998). We start with 3D Green's function for homogeneous media

$$g(\mathbf{x}, \mathbf{x}_s, \omega) = \frac{e^{i\omega\tau_s}}{4\pi r_s}, \tag{2.69}$$

where r_s is the radius from the source to the image point.

Plug in this Green's function for both source and receiver wavefields into the formula of gradient. For a single source, we obtain

$$\gamma(\mathbf{x}) = \int d\mathbf{x}_g \int d\omega \; Re[\frac{1}{16\pi^2} F(\omega) \frac{1}{r_s r_g} \omega^2 \delta u^*(\mathbf{x}_g, \mathbf{x}_s, \omega) e^{i\omega(\tau_s+\tau_g)}]. \tag{2.70}$$

Letting

$$\delta u'(\mathbf{x}_g, \mathbf{x}_s, \omega) = F(\omega)\delta u(\mathbf{x}_g, \mathbf{x}_s, \omega), \tag{2.71}$$

we get

$$\gamma(\mathbf{x}) = \int d\mathbf{x}_g \frac{1}{r_s r_g} \frac{1}{16\pi^2} \int d\omega \; \omega^2 \delta u'^*(\mathbf{x}_g, \mathbf{x}_s, \omega) e^{i\omega(\tau_s+\tau_g)}. \tag{2.72}$$

Taking the conjugation to the whole integrand, we get

$$\gamma(\mathbf{x}) = -\frac{1}{8\pi} \int d\mathbf{x}_g \frac{1}{r_s r_g} \frac{1}{2\pi} \int d\omega \; [(-i\omega)(-i\omega)\delta u(\mathbf{x}_g, \mathbf{x}_s, \omega) e^{-i\omega(\tau_s+\tau_g)}]^*. \tag{2.73}$$

Transforming equation 2.73 to the time domain, we get

$$\gamma(\mathbf{x}) = -\frac{1}{8\pi} \int d\mathbf{x}_g \frac{1}{r_s r_g} \ddot{\delta U'}(\mathbf{x}_g, \mathbf{x}_s, \tau_s + \tau_g), \tag{2.74}$$

where

$$\delta U'(\mathbf{x}_g, \mathbf{x}_s, t) = F(t) * \delta U(\mathbf{x}_g, \mathbf{x}_s, t). \tag{2.75}$$

We compare now the result we obtained in equation 2.74 to the Kirchhoff migration formula adapted from (Yilmaz, 2001, p. 1346)

$$I(\mathbf{x}) = \frac{1}{4\pi} \int d\mathbf{x}_g \frac{cos\theta}{v r_s} \dot{U}_{obs}(\mathbf{x}_g, \mathbf{x}_s, \tau_s + \tau_g), \tag{2.76}$$

where $I(\mathbf{x})$ is the output image and $cos\theta$ is the obliquity factor.

We note here that that waveform inversion uses the second time derivative of the data residuals $\delta\ddot{U}'(\mathbf{x}_g, \mathbf{x}_s, \tau_s + \tau_g)$ whereas in Kirchhoff migration the first time derivative of the observed data themselves is used $\dot{U}_{obs}(\mathbf{x}_g, \mathbf{x}_s, \tau_s + \tau_g)$. The second difference is that the correction factor used, $\frac{1}{r_s r_g}$ is used for waveform inversion whereas that used for Kirchhoff migration is $\frac{cos\theta}{v r_s}$. As pointed out by Tarantola (1984b), those differences are due to the fact the waveform inversion is trying to invert for a different quantity, namely the velocity. However, the kinematics of the two methods are the same since both spread the data over the isochrones of time $\tau_s + \tau_g$.

2.4.6 Amplitude preserving migration/inversion

Amplitude preserving migration (APM) or true amplitude migration is a migration process that produces a migrated image which preservers the relative amplitudes. The migration method by Bleistein et al. (2001) produces an image whose amplitudes honors the transport equation. In this section we will try to compare it to waveform inversion.

Bleistein et al. (2001) use an approximation to the Green's functions, using the WKBJ method, to arrive to their inversion formula. Waveform inversion, on the other hand, uses the more accurate numerical methods, such as finite difference, to calculate the Green's functions. Although the two methods are intended to evaluate two different quantities,

velocity for waveform inversion and band-limited reflectivity for APM migration, they are similar. The formula adapted from (Bleistein et al., 2001, p. 247)is

$$\beta_1(\mathbf{x}) = \frac{1}{\pi c}\int d\omega \int d\mathbf{x_s} \int d\mathbf{x_g}\,(\frac{r_s}{r_g^2})\,(i\omega)\,e^{-i\omega[r_s+r_g]/c}u_{obs}(\mathbf{x_g},\mathbf{x_s},\omega), \tag{2.77}$$

where β_1 is the APM image with , $r_s$and $r_g$are the distances between the source and image point and the receiver and image point, respectively. It can also be written as

$$\beta_1(\mathbf{x}) = \frac{1}{\pi c}\int d\omega \int d\mathbf{x_s} \int d\mathbf{x_g}\,(\frac{r_s}{r_g^2})\,(i\omega)\,e^{-i\omega(\tau_s+\tau_g)}u_{obs}(\mathbf{x_g},\mathbf{x_s},\omega). \tag{2.78}$$

Notice that the main difference between equation 2.78 and equations 2.76 and 2.74 is roughly the weighting factor but they are kinematically equivalent.

2.4.7 Waveform inversion and traveltime tomography

Traveltime tomography and waveform inversion are closely related. The first to point this out is Woodward (1992). Traveltime tomography is based on the integral equation which is

$$\delta t(\mathbf{x}_g, \mathbf{x}_s, \omega) = \int d^3x \mathcal{D}_{\mathbf{ray}}(\mathbf{x}, \mathbf{x_g}, \mathbf{x_r})\frac{1}{\delta v(\mathbf{x})}, \tag{2.79}$$

where $\delta t(\mathbf{x}_g, \mathbf{x}_s, \omega)$ is the traveltime residual and $\mathcal{D}_{\mathbf{ray}}(\mathbf{x}, \mathbf{x_g}, \mathbf{x_r})$ is raypath. Waveform inversion on the other hand is based on the first order Born approximation. For comparison we state here again the Born integral formula

$$\delta u(\mathbf{x}_g, \mathbf{x}_s, \omega) = \omega^2 F(\omega)\int_V dx^3 g(\mathbf{x}_g, \mathbf{y}, \omega)g(y, \mathbf{x}_s, \omega)\delta m(\mathbf{y}) \tag{2.80}$$

Let $\mathcal{D}_{wave} = g(\mathbf{x}_g, \mathbf{y}, \omega)g(y, \mathbf{x}_s, \omega)$. Then the integral equation becomes

$$\delta u(\mathbf{x}_g, \mathbf{x}_s, \omega) = \omega^2 F(\omega)\int_V dx^3 \mathcal{D}_{\mathbf{wave}}(\mathbf{x}, \mathbf{x_g}, \mathbf{x_s})\delta m(\mathbf{x}), \tag{2.81}$$

$\mathcal{D}_{wave}$ is analogous to the raypath $\mathcal{D}_{ray}(\mathbf{x}, \mathbf{x}_g, \mathbf{x}_r)$ but it is based on waves instead of rays. Hence Woodward (1992) referred to it as the wavepath. Therefore, traveltime

tomography is based on back projecting the the time residuals over the raypaths and waveform inversion is based on back projecting the data residuals along the wavepaths (Woodward, 1992).

2.5 Nonlinearity

Since waveform inversion is highly non-linear, certain steps need to be taken in order to overcome converging into a local minimum. Several strategies have been developed and we will discuss them in this section.

2.5.1 Multiscale strategy

The mutliscale strategy is simply based on the idea of updating the model first with the large features and progressively adding more details or the small features. The first paper to show the effectiveness of such strategy on waveform inversion is that of Bunks et al. (1995). They used multiple grids as as the name suggests. Early iterations of inversion use sparse grid and later iterations use progressively finer and finer grid. An equivalent strategy is to invert the low frequencies of the data first and progressively higher and higher frequencies (Sirgue and Pratt, 2004a).

2.5.2 Damping late arrivals

The late arrivals are often dominated by reflections and multiple arrivals. Damping or windowing out the late arrivals yields an objective function that is much better behaved, i.e. smoother and has fewer local minima as Sheng et al. (2006) show. Apparently, the first paper to advocate such a strategy was Pratt and Worthington (1988).

2.5.3 Preconditioning the updates/gradient

As we have stated in section 2.5.1, preconditioning the gradient of the misfit function is a very important step since the gradient is operated upon by a roughening operator. It is therefore, necessary to stabilize the inversion. The simplest method of preconditioning the gradient is simple convolutional smoothing. However, more sophisticated methods have been developed and tested, such as using the inverse of the approximate Hessian Sheng et al. (2006), dynamic smoothing of Nemeth et al. (1997) and filtering in the wavenumber domain of Sirgue (2003). The method that is implemented in the software package of Pratt and coworkers is wavenumber domain filtering .

2.6 Conclusion

We derived in this chapter the important formulas in waveform inversion. We discussed the forward problem and derived the formula for the gradient of the misfit function of waveform inversion. We also compared waveform inversion with Kirchhoff migration, true amplitude-preserving migration and traveltime tomography. We also discussed the properties of waveform inversion, the necessity of preconditioning and the rest of strategies used to mitigate the highly non-linear nature waveform inversion.

This page is intentionally left blank.

Chapter 3

Waveform inversion and local heterogeneity

3.1 Introduction

The near surface is often characterized by strong heterogeneity. Unconsolidated sediments, gravels and hard rocks, to name a few causes, can all be found side by side in the near surface. This chapter examines the effect of such abrupt strong variations in velocities on waveform inversion and whether certain strategies are needed. For this purpose we created two velocity models: one with those variations, and one without. We inverted both of them using the sequential strategy. We used different initial models and different starting frequencies. We also examined the efficient strategy of Sirgue and Pratt (2004a).

3.2 Modeling

Our objective is to examine the effect of heterogeneity on waveform inversion. For this purpose, we created two velocity models. The first velocity model is shown in Figure 3.1(c), hereafter referred to as model 1. The prominent characteristics of this model are the sinusoidal feature and the velocity gradient within the layers. Although this model might look simple, it is not. A quick look at the synthetic data generated for this model as shown in Figure 3.3, clearly demonstrates this point. The diffractions and reflections and scattering in general are very predominant.

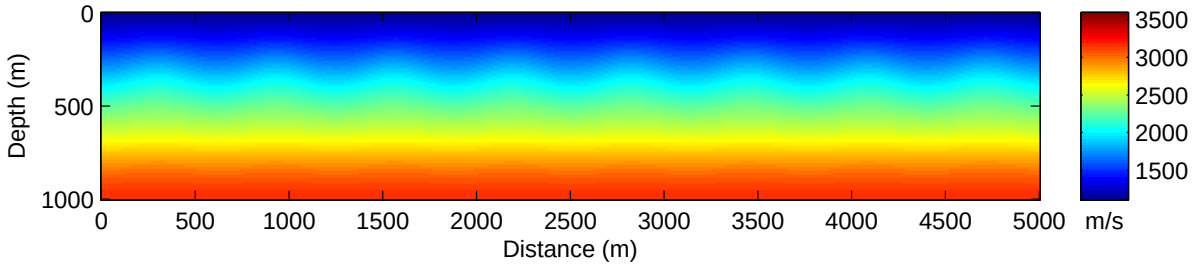
We include the sinusoidal feature in the model because such a feature causes long wavelength statics problems, which are not resolved using residual statics (Yilmaz, 2001; Cox, 1999). This model is somewhat similar to that used by (Sheng et al., 2006).

The second velocity model 3.2(c), hereafter referred as model 2, is the same as model 1

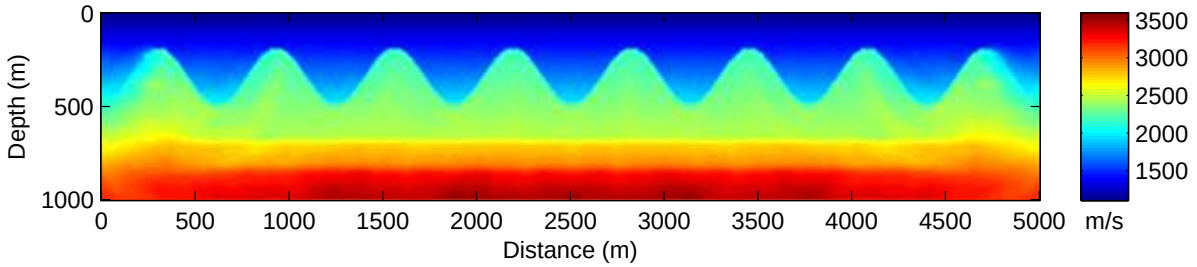
except that some random velocity variations are added to it. The variations are normally distributed with a mean of 500 m/s and a standard deviation of 200 m/s. The model is then smoothed by a 20 m x 20 m smoother since our goal is not to drastically make model 2 different from model 1, but rather to make the large features similar. Model 2 exhibits some abrupt velocity variations, a general characteristic of the near surface.

The size of the models is 500 x 100 grid points in x and z respectively, where each grid point corresponds to 10 m. Synthetic data are then acquired over these models. The receiver interval is 20 m and the source interval is 25 m. The location of the first source is 115 m from the left side of the model and the location of the first receiver is 110 m. The depth of the sources and receivers is 11 grid points below the top of the model since the sponge boundary conditions in the corners need at least 10 points to be effective. The total number of shots is 191 and the number of traces per shot is 240. We use frequency-domain finite-difference modeling, to generate the synthetic data. The top boundary is a free surface, whereas the rest are absorbing. A sponge boundary condition is used for the corners. A sample time-domain shot record whose location is in the middle of model 1 is shown in Figure 3.3(a).

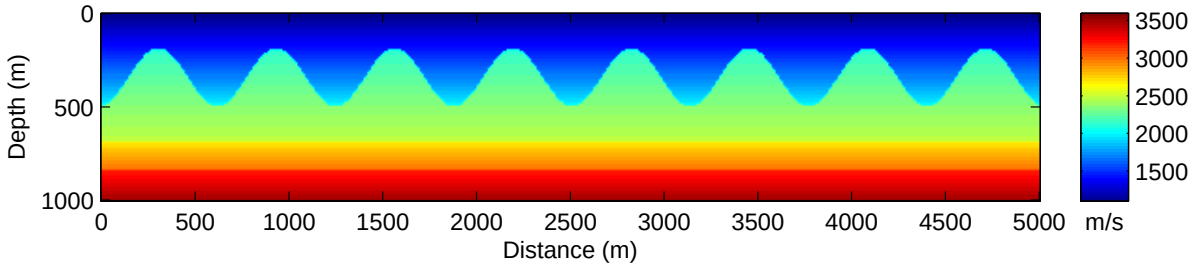
Since the modeling is carried-out in the frequency domain, the data generated are obviously in the frequency domain, as well. Therefore, we show the data in frequency-domain. Figures 3.4 and 3.5 show the real part of the frequency domain panels plotted as surfaces, as a function of source and receiver locations, for two frequencies. This representation can be useful in analyzing the data. The effect of the sinusoidal feature of the model is clear for the higher frequency panel in Figure 3.4, but for the lower frequency it is smoother. In other words, the low frequencies of the data are obviously sensitive to the large features of the model but not its small details. This explains why the low frequencies are needed for inversion since they are more sensitive to the low wavenumber component of the model (Sirgue, 2003). Since model 2 is more complex than model 1,



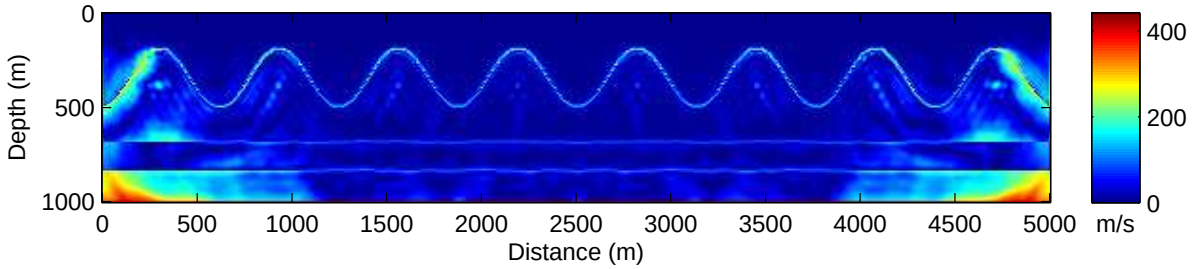
(a) Initial Model



(b) Estimated Model

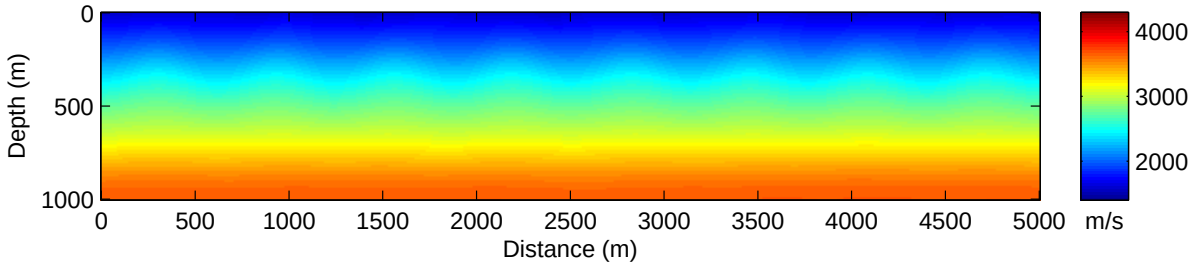


(c) True Model

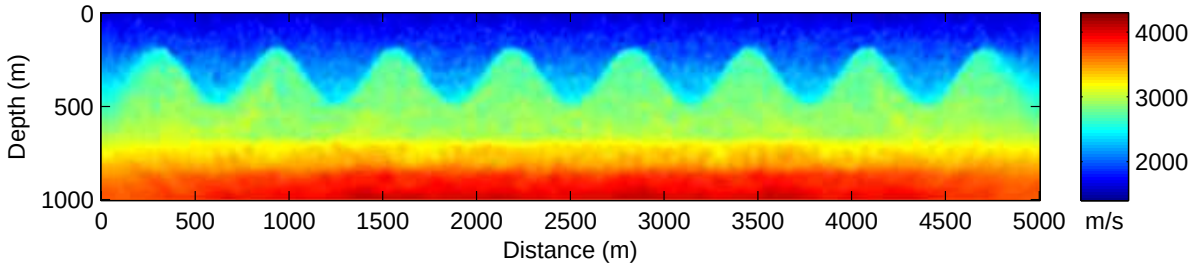


(d) Absolute value of the difference between a and b

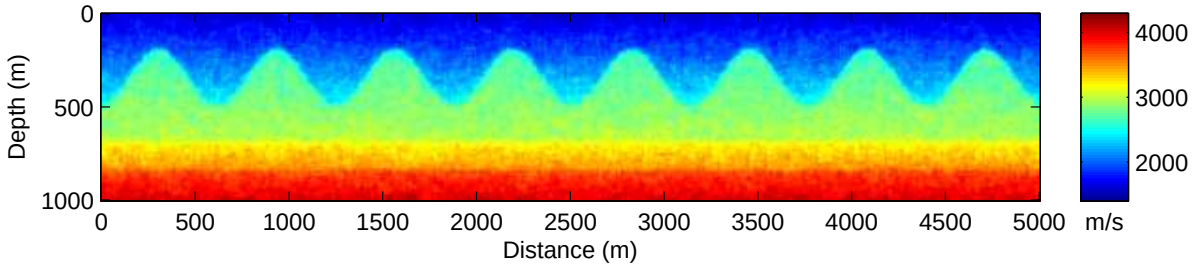
Figure 3.1: Velocity models associated with model 1. The inversion scheme is referred to, in the text, as Part 1 of the inversion. The initial model is the true model smoothed by 300 x 300 m smoother. The sequential inversion strategy is used. The frequencies inverted are from 2-24.8 Hz with 0.2Hz frequency interval.



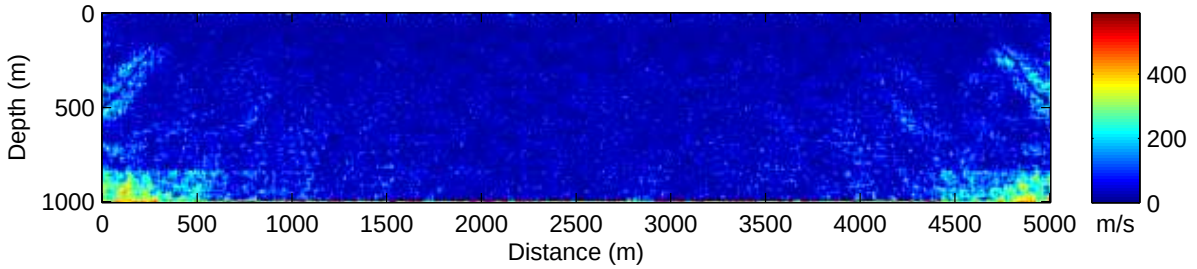
(a) Initial Model



(b) Estimated Model



(c) True Model



(d) Absolute value of the difference between a and b

Figure 3.2: Velocity models associated with model 2. The inversion scheme is referred to, in the text, as Part 2 of the inversion. The initial model is the true model smoothed by 300 x 300 m smoother. The sequential inversion strategy is used. The frequencies inverted are from 2-24.8 Hz with 0.2Hz frequency interval.

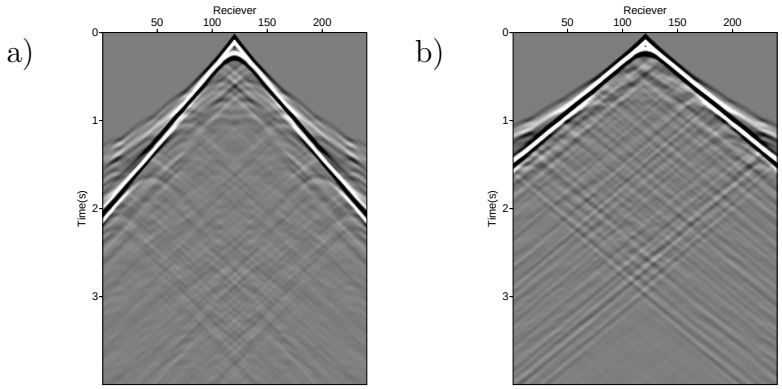


Figure 3.3: Sample shot records, a) for model 1 and b) for model 2.

the effect of the sinusoidal feature is less clear as shown in Figure 3.5; but the lower frequency panel is still smoother.

3.3 Inversion

3.3.1 Sequential strategy

Part1: higher resolution initial model and low starting frequency

Since the result of waveform inversion strongly depends on the lowest frequency used and the resolution of the initial model, we start the inversion with a higher resolution initial model and lower starting frequency. We use an initial model which is the true model smoothed by a 300 m x 300 m smoother. The starting frequency used in inversion is 2 Hz. The initial model for model 1 is shown in Figure 3.1(a), whereas that for model 2 is shown in Figure 3.2(a).

We inverted the data for both models, using the multiscale strategy, where low frequencies are inverted first and then higher frequencies. In other words, the model is updated first using the low frequencies and then we progressively invert higher and higher frequencies.The highest frequency component of the data used is 24.8 Hz. The number of sequential iterations per frequency is 5, and the frequency interval is 0.2 Hz. The

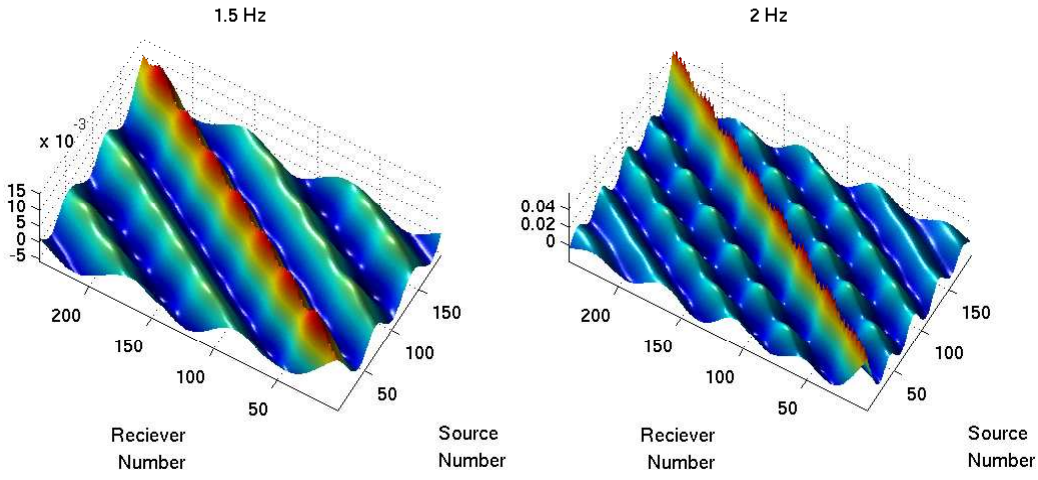


Figure 3.4: Frequency domain panels (real part) for seismic data of model 1.

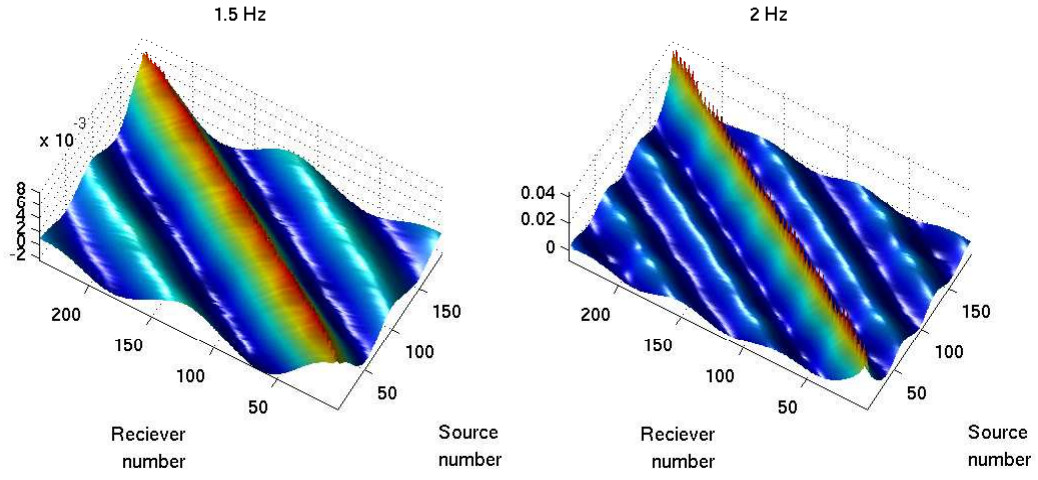


Figure 3.5: Frequency domain panels (real part) for seismic data of model 2.

gradient of the misfit function was preconditioned by a low-pass filter in the wavenumber domain (Sirgue, 2003), as we discussed in Chapter 2.

The final inversion result of model 1 is shown in Figure 3.1(b). Figure 3.1(d) shows the absolute value of the difference between the true and the estimated models. The result is very well-resolved in most areas. The exact interfaces are not exactly resolved since they might require even higher frequencies. The right and left edges of the model are not resolved to the same degree as the middle part of the model, and this is due to the finite aperture of the source-receiver coverage. The same analysis applies to the result of inverting model 2, as shown in Figure 3.2(b) and 3.2(d). Because the interfaces are not as well defined as in model 1, the inversion result of model 2 is even more resolved at the interfaces. The small velocity variations in model 2, which are beyond the resolution, are not completely recovered as shown in Figure 3.2(d).

Figures 3.6(a) and 3.6(b) show the L2-norm of the difference between the true and the estimated models for each depth. The red curve excludes the left and the right-hand edges of the models since the error is higher in these areas. The blue curve includes all of them. Both Figures 3.6(a) and 3.6(b) show a general trend. In Figure 3.6(a), the error is particularly higher within the sinusoidal structure since it is more complex than the flat layers below. On the other hand, Figure 3.6(b) does not exhibit this characteristic clearly, since complexity is randomly distributed in the model. Note that the error is higher on the top part of Figure 3.6(b), and this is due to the fact that the receivers are at depth in the model.

Part 2: Smoother initial model and higher starting frequency

Now we challenge the inversion method by using a smoother initial model and a higher starting frequency. The initial model is the true model smoothed by a 500 m x 500 m smoother. The lowest frequency inverted is 5 Hz. Both of those values were used by

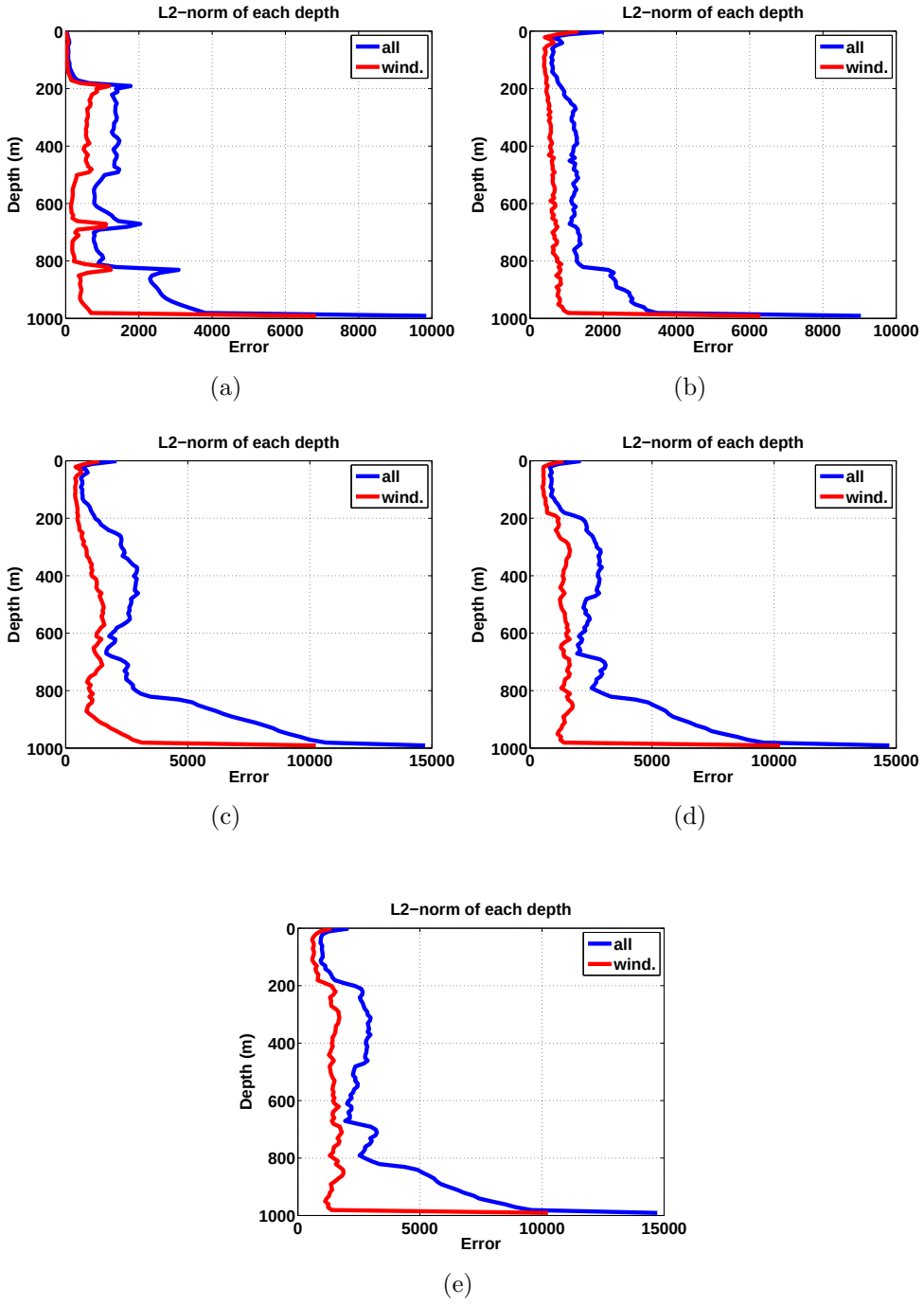


Figure 3.6: L2 norm of the difference between the true and the estimate models for each depth. a)-e) are associated with Figures 3.1, 3.2, 3.7, 3.8 and 3.9.

Ravaut et al. (2004). The rest of the inversion parameters are identical to those of Part 1 from the previous section.

Figure 3.7(b) shows the result of the inversion. Although the result is still well resolved, it is not as accurate as that of Part 1. The L2 norm of the difference of each depth is also higher for this result than from Part 2, as shown in Figures 3.6(b) and 3.6(c). This result clearly shows that when using this inversion scheme, the inversion result is less accurate if higher frequencies are used and, as well as, a less resolved initial model.

3.3.2 Efficient strategy

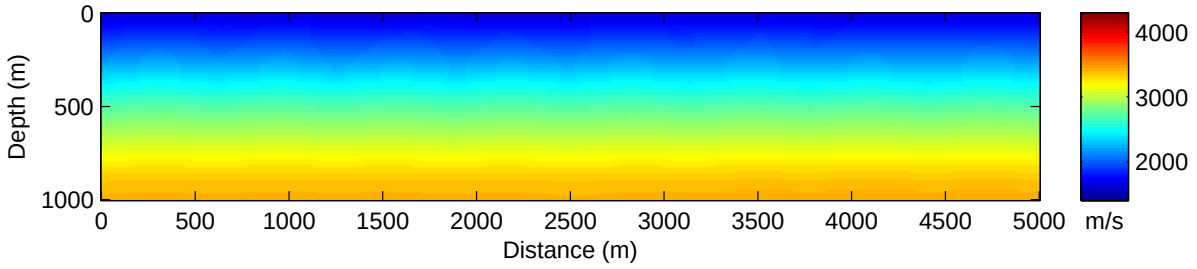
Now we apply the efficient strategy of Sirgue and Pratt (2004b), as we have discussed in Chapter 2. The strategy provides a way to take advantage of the redundancy of information and hence select as few frequencies as possible for inversion. The first step is to calculate α_{min}, which is equivalent to the inverse NMO (normal-moveout) stretch factor as (Sirgue and Pratt, 2004b), as follows:

$$\alpha_{min} = \frac{1}{\sqrt{1 + (\frac{h_{max}}{z})^2}}, \tag{3.1}$$

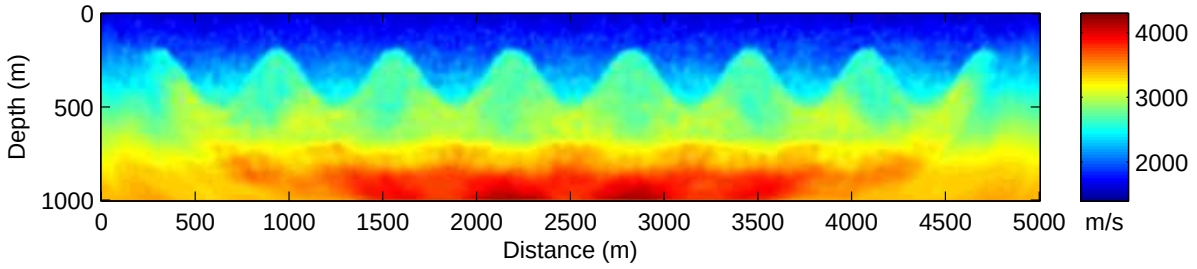
where h_{max} is half the offset, and z is the depth. Then we can calculate the frequencies from a starting frequency as follows:

$$f_{n+1} = \frac{f_n}{\alpha_{min}} \tag{3.2}$$

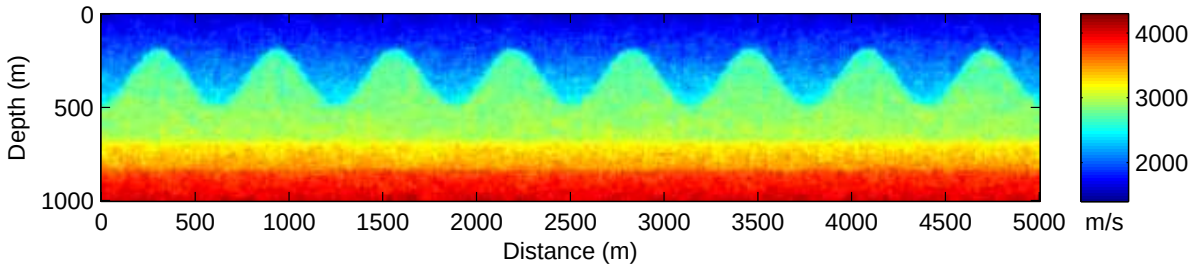
The depth of the model is 1000 m, which is z. Since the width of the interior of the model that is properly imaged is around 3000 m in x, we can set h_{max} to 1500 m. Therefore, starting from 5 Hz, the strategy then suggests using 9 Hz, 16.3 Hz and 29.3 Hz. The highest frequency we have available is 24.8 Hz. The inversion was carried out using parameters identical to the previous examples, except that the number of sequential



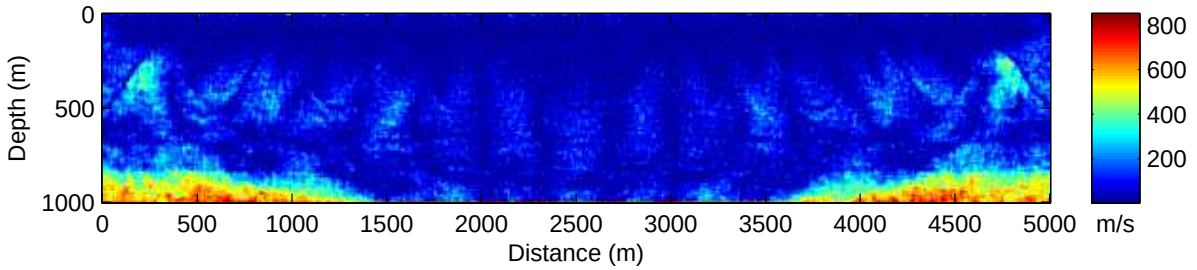
(a) Initial Model



(b) Estimated Model



(c) True Model



(d) Absolute value of the difference between a and b

Figure 3.7: Inversion of dataset 2 using a smoother initial velocity model (500 m x 500 m smoother dimensions) and a higher starting frequency, 5Hz. The efficient strategy is not used.

iterations per frequency was increased from 5 to 30, in order to further minimize the misfit function (Brenders and Pratt, 2007b).

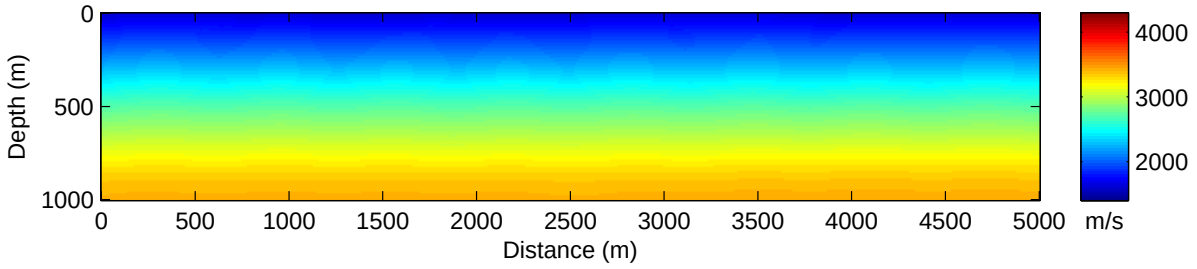
Figure 3.8 shows the result using 5, 9 and 16.3 Hz. Figure 3.9 shows the result using 5, 9, 16.3 as well as the highest available frequency, 24.8 Hz. Notice that minimal improvement was achieved by inverting the highest available frequency, 24.8 Hz.

Comparing the efficient strategy result in Figure 3.8 and the result in Figure 3.7, we can observe that they are comparable, although the very shallow part of the model seems to be better resolved in the case of the non-efficient method. The very deep part of the model on the other hand seems to be better imaged using the efficient method as shown in Figures 3.6(c) and 3.6(d).

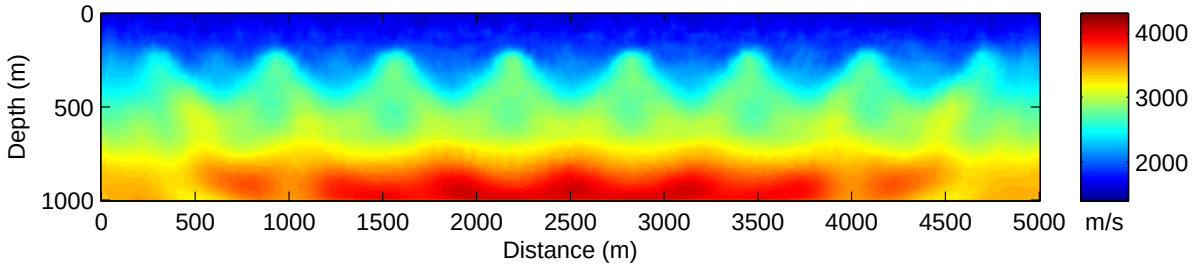
Since the number of iterations used per frequency is 30, the total number of iterations for the efficient method using three frequencies is 90 iterations. On the other hand, the total number of iterations used in the non-efficient method is 500 iterations. In this case, the efficient method costs almost 20% of the cost of the non-efficient method.

3.4 Conclusion and discussion

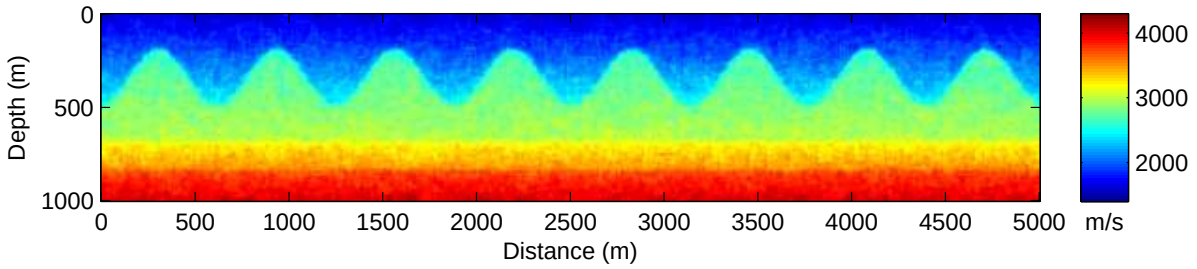
Waveform inversion provides an effective and efficient way of imaging the near surface. We showed that even in the presence of locally strong variations of velocity, waveform inversion is capable of resolving such models. Local heterogeneity has minimal effect on waveform inversion if an accurate initial model is used and low frequencies are inverted. Waveform inversion provided models with very high resolution regardless of the heterogeneity of the model. The efficient strategy was able to recover the velocity model with comparable resolution of that of the non-efficient method, but with about 80% decresase in the computational cost.



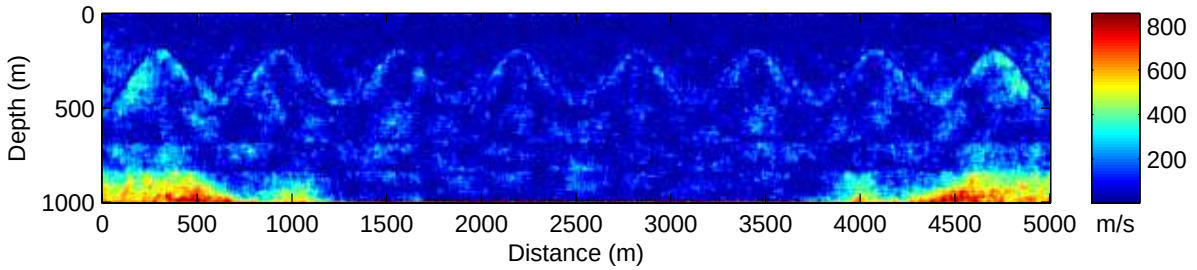
(a) Initial Model



(b) Estimated Model

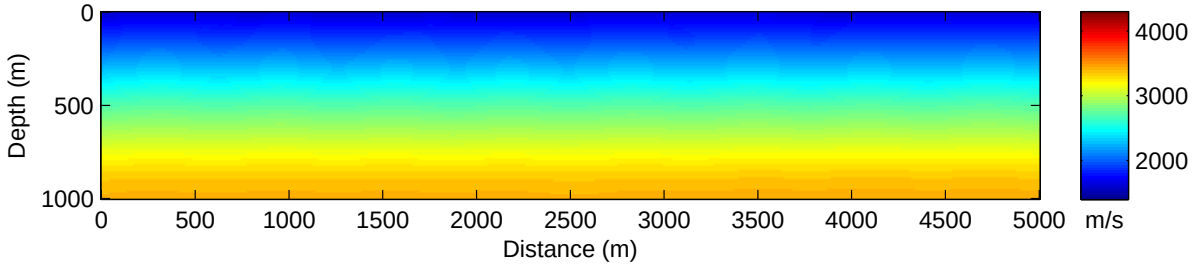


(c) True Model

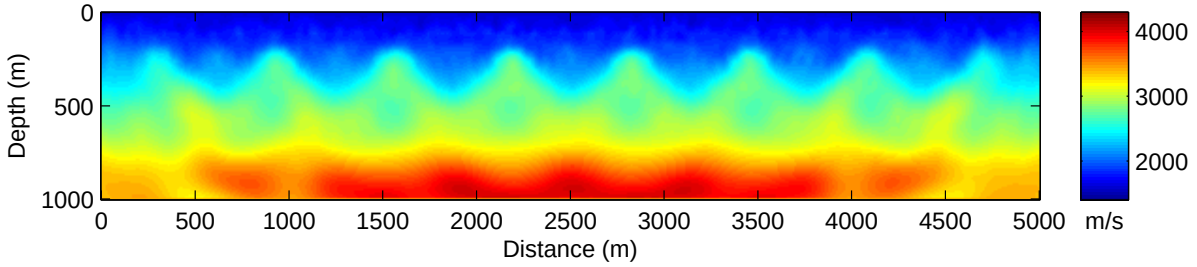


(d) Absolute value of the difference between a and b

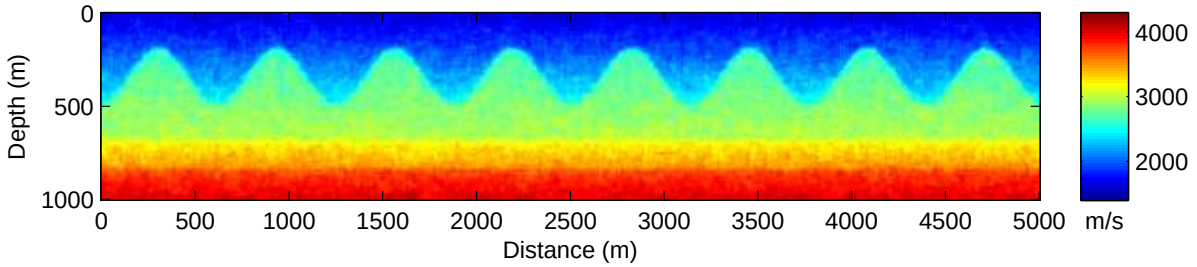
Figure 3.8: Inversion of dataset 2 using a smoother initial velocity model (500 m x 500 m smoother dimensions) and a higher starting frequency, 5Hz. The efficient strategy is used.



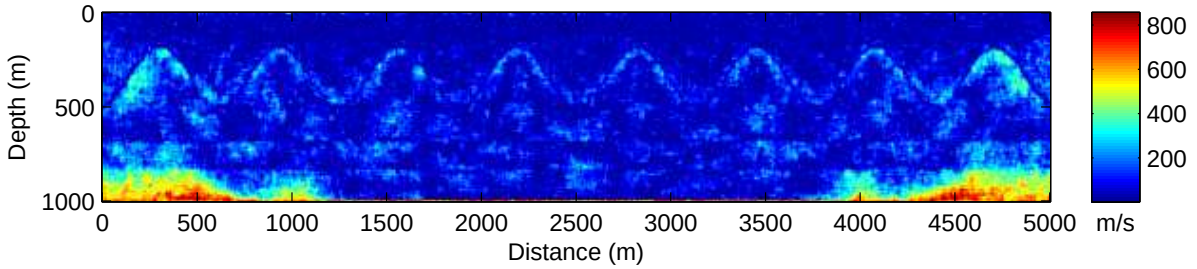
(a) Initial Model



(b) Estimated Model



(c) True Model



(d) Absolute Value of the Difference Between a and b

Figure 3.9: Inversion of dataset 2 using a smoother initial velocity model (500 m x 500 m smoother dimensions) and a higher starting frequency, 5Hz. The efficient strategy is used plus the highest available frequency, 24.8 Hz.

This page is intentionally left blank.

Chapter 4

Waveform inversion and complex near surface

As we have shown in chapter 3, the benefits applying waveform inversion to areas of complex targets are clear. Our focus, however, is on areas of complex near surface. For this purpose, conventional methods like traveltime tomography and refraction statics are often applied. Although those methods are robust, they are only applicable to relatively simple near surface models. Complex near surface models need more rigorous solutions. In this chapter, we apply waveform inversion to an area of very challenging realistic near surface geology. The complex near surface of the model makes it difficult to image a low relief structure at depth. The goal of the inversion is to recover the near surface compressional velocity and to detect some indication of the of the low relief structure. We inverted the data using both the sequential and the efficient strategies. We compare the statics provided by the estimated models.

4.1 Introduction

Many methods exist for inversion of the near surface velocities. Refractions statics, traveltime tomography and others are robust and efficient methods for this purpose (Yilmaz, 2001, p. 370). However, in areas with complex near surface geology, conventional methods are challenged. We use frequency-domain waveform inversion to go beyond the traveltimes, and extract the compressional velocity information of the near surface from the waveforms directly.

Waveform inversion has been applied to synthetic and real data (e.g. Sheng et al., 2006). It also has been applied, in one case study, for the purpose of extracting near

surface compressional velocities (Sheng et al., 2006). However, the areas that most need waveform inversion are the ones for which conventional methods fail. In this study, we apply waveform inversion to a model of complex near surface geology. The model bears some near surface challenges that obscure the subtle feature buried underneath, the low relief anticlinal structure.

We first describe the model and its complexity. Then, we generate synthetic data using the model and invert them. We first invert the data using a sequential scheme. Then, we test the efficient strategy of Sirgue and Pratt (2004a). Finally, near-surface vertical traveltimes, or statics, are calculated from the estimated models and compared with the true statics.

4.2 Model description and Synthetic data

4.2.1 The model

Figure 4.1 shows the model in detail. Alkhalifah and Bagaini (2006) describe the model fully. We give a brief description here. The model bears some of the near surface challenges observed in a complex near surface desert environment such as Saudi Arabia. The sand dunes with their relatively low velocity cover some parts of the surface. Some of them can be over 150 meters high. The wadi in the middle contains some relatively low velocity unconsolidated sediments. Below those features is a thin low velocity zone (LVZ) that is often very difficult to detect and properly correct for.

Below the LVZ are other low velocity anomalies, which model the collapse features within the Rus formation. The collapses result from the dissolution of carbonates within this formation. Below this complex burden is a structure with low relief. The complex near surface and the subtle relief of the structure makes imaging and delineating the anticlinal structure a difficult task.

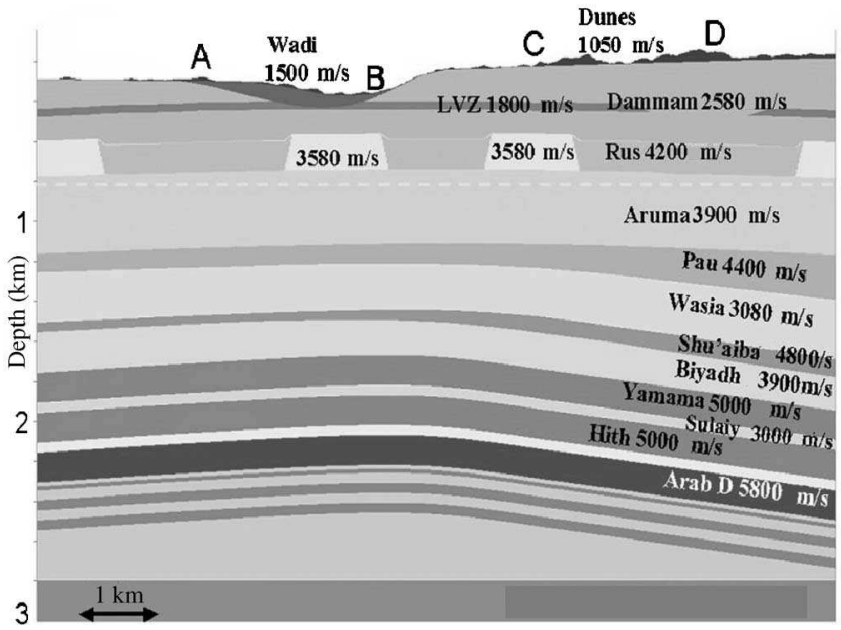


Figure 4.1: The velocity model used in this study, from Alkhalifah and Bagaini (2006).

The elevation profile of the model is shown in Figure 4.2. The elevation can vary by over 200 meters. The picture in Figure 4.3 shows some of the complexity of the near surface in Saudi Arabia and the topographical challenges.

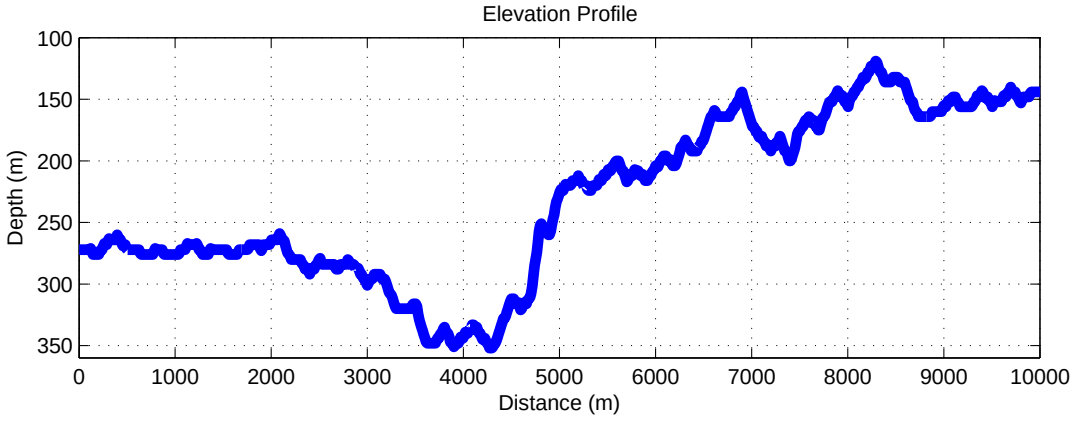


Figure 4.2: The elevation profile of the model.

Waveform inversion is based on inverting mainly transmissions (Pratt, 2008; Sheng et al., 2006). In order to determine the maximum depth of the model that can be

Figure 4.3: An example area that shows how extreme the near surface and the topography can be in Saudi Arabia. Picture from Wikipedia (2009).

reliably inverted, we model diving waves. Figure 4.4 shows, the diving waves modeled in a smoothed version of the model. We can see that the diving waves could reach about 2 km of depth. Therefore we truncate the model to this depth, since there is no point of including the rest of the model that may not be recovered reliably.

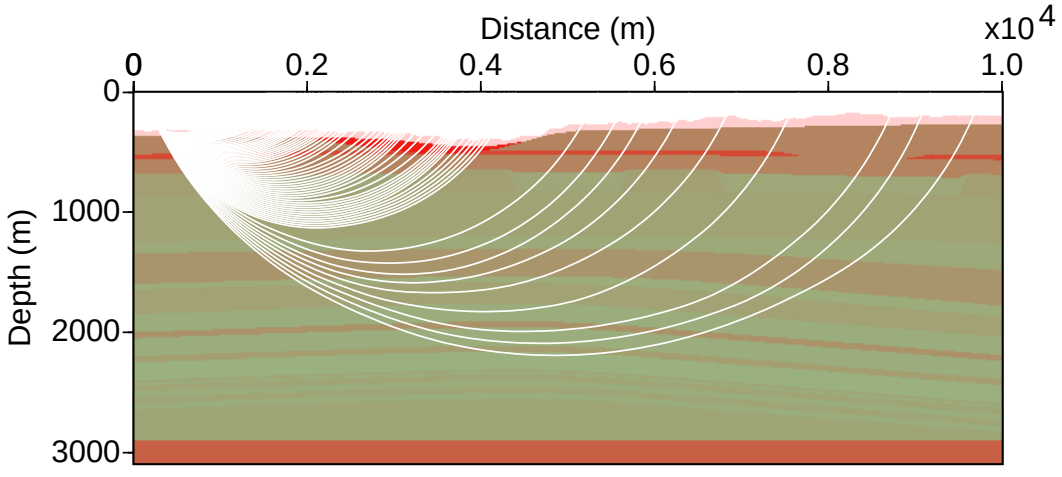


Figure 4.4: Diving rays traced in a smoothed version of the model.

4.2.2 Synthetic data

The original synthetic data has a fairly limited offset. To image the full near surface and some reflectors at depth, we have generated the data with a maximum offset of about 9500 m. The receiver interval, for the dataset is 25 m and the source interval is 35 m. The entire spread is alive for this experiment. The source wavelet is a Keuper wavelet with 5.5 Hz dominant frequency, similar to that used by Brenders and Pratt (2007b). Figure 4.5(a) shows three sample shot records.

The frequency-domain finite-difference software package does not allow non-flat free surface. Including the free surface effects would require modification to the already well-tested and robust package, a non-trivial task. After all, Brenders and Pratt (2007b,a) have successfully inverted viscoelastic data without including the free surface effects. For those reasons, we do not use a free surface for both modeling and inversion. In fact, Brenders and Pratt (2007b) have successfully inverted the viscoelastic data using pressure formulation, whereas the data are in particle velocity.

4.3 Preliminary inversion

It is customary to show the final result of the inversion without discussing the intermediate ones, or what lead us to the final result. Here, we avoid this approach. We show how we arrived to the final result, by discussing the easier tests we conducted.

Our convention of organizing the results, just like in chapter 3, will be to show the initial model on the top, then the estimated model, the true model and finally the absolute value of the difference between the true model and the estimated one. We will follow this convention so that the reader could compare the models on the same page rather than on completely separate pages, although this convention might lead to repeating some the figures. The 1D profiles will immediately follow the figures containing the full models,

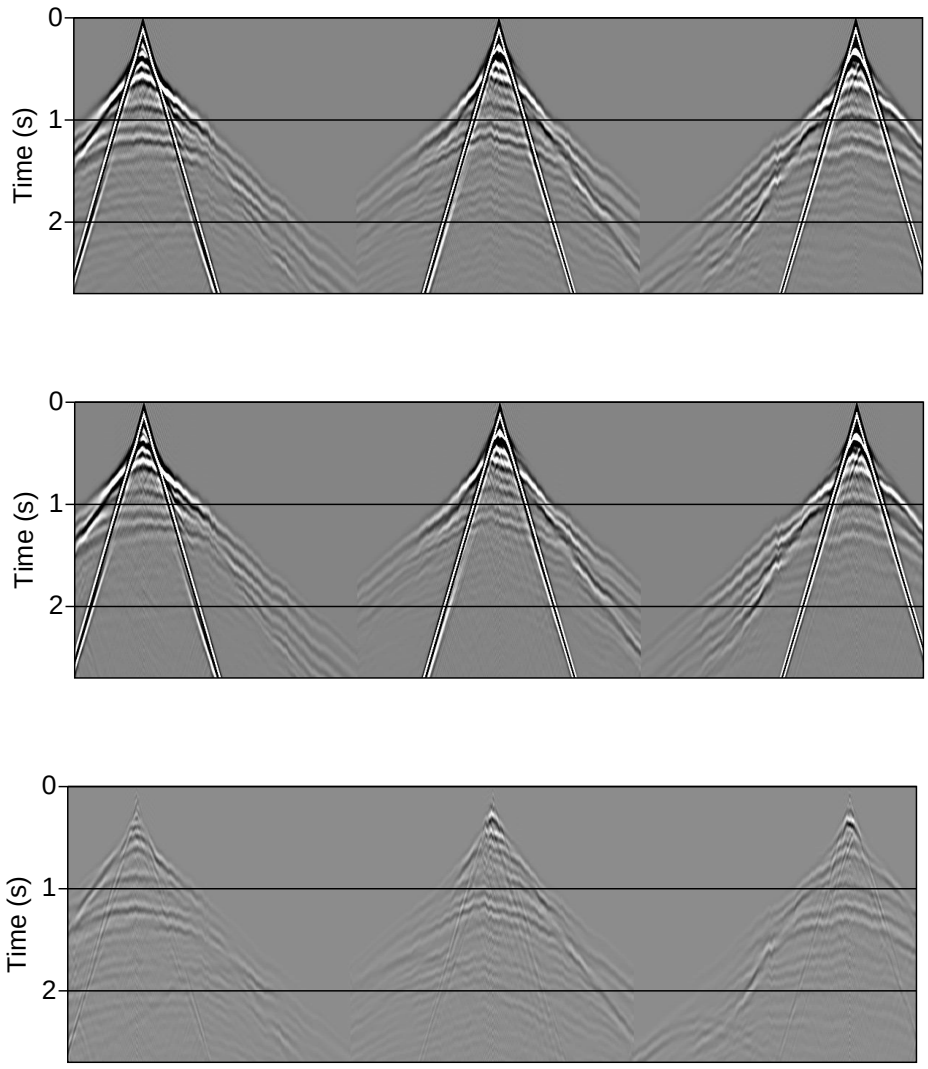


Figure 4.5: Sample shots for a) the true model, b) the estimated model from the sequential experiment and c) the difference between them.

so that more quantitative comparison can be made.

4.3.1 Sequential Strategy

We start the inversion by using an initial model that is relatively close to the true model. The initial model is the true model smoothed by a 240 m x 240 m smoother. Just as in Chapter 2, we use the Seismic Un*x program *smooth2* with a parameters of 20 x 20 as the smoothing factors. Given that the grid size is 12 m, that is almost equivalent to using a smoother with dimensions 240 m x 240 m. The initial model in this case simulates one that can be obtained for a mature field. The plethora of data collected over decades for such fields makes building such an initial model possible. Well logs, borehole data, check shots, VSP, crosswell data, traveltime tomography, refraction statics, migration velocity analysis and other methods are commonly used in the industry for building such models as detailed in the book by Yilmaz (2001). As we will see, what waveform inversion will obtain a model with substantially higher resolution.

We inverted the frequencies from 2-14.8 Hz sequentially with 0.2 Hz interval. In other words, we inverted 2 Hz by itself and updated the initial model. Then we inverted 2.2 Hz and updated the model from the previous frequency, and so on and so forth. This strategy is essentially equivalent to using multiple scales or multigrid, and therefore, it is often refereed to as the multiscale strategy (Sirgue and Pratt, 2004a; Pratt, 2008; Bunks et al., 1995). Since we are using a gradient based method, a number of iterations needs to be performed in order for the method to converge. We perform 5 iterations per frequency just as in Brenders and Pratt (2007b).

The result of the inversion is shown in Figure 4.6(b). The inversion has much higher resolution than the initial model. More importantly, it accurately resembles the true model. The important and difficult features are accurately resolved. The Wadi, the thin low velocity zone and the low velocity zones within the Rus formation are all in the

inversion result. As Figure 4.6(d) shows, the error of the result is less than about +/- 400 m/s. It is actually much better than that in most places as shown in the 1D velocity profiles in Figure 4.7.

4.3.2 Efficient Strategy

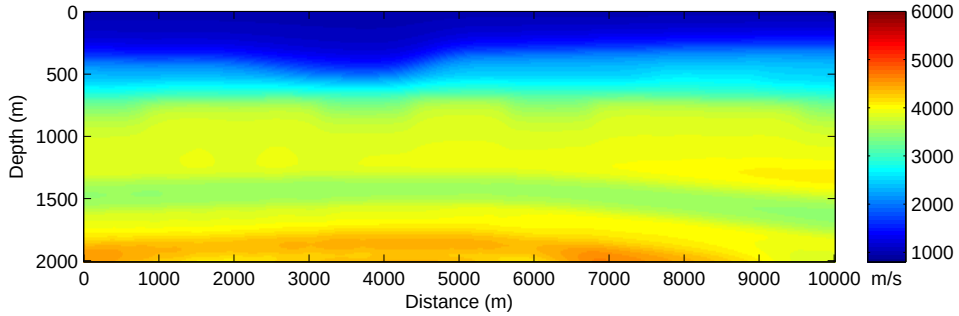
Now, we will put the efficient strategy to test. Using the same method mentioned in Chapter 2 and 3, the inversion strategy suggests using 2, 4.5, 10 and 22.4 Hz only. The inversion result is shown in Figures 4.8(b) and the 1D profiles in Figure 4.9. The inversion parameters used are identical to those used in the previous experiment. Although the result is more accurate than the initial model, it is as accurate as that resulting from the sequential strategy. This does not imply, however, that more accurate models cannot be obtained using the efficient strategy. Quite the contrary. More accurate models can indeed be obtained. However, more aggressive preconditioning of the updates/gradients needs to be performed. This makes the results more subjective and human involved. But our goal is lessen the human intervention as much as possible.

4.4 Final inversion

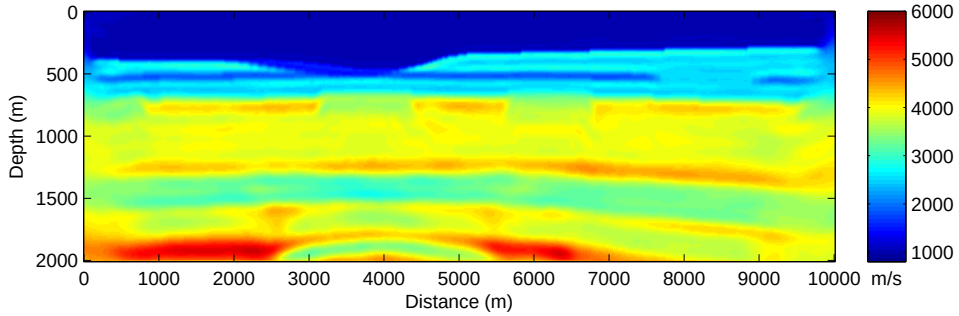
4.4.1 Sequential Strategy

We inverted for the model using an initial model that is a much more smooth version of the true model. The smoother used is about 480 m x 480 m. We inverted sequentially each frequency from 2-14.8 Hz, with an interval of 0.2 Hz. We used 5 iterations per frequency. The preconditioning used is based on filtering the gradient of the misfit function in the wavenumber domain.

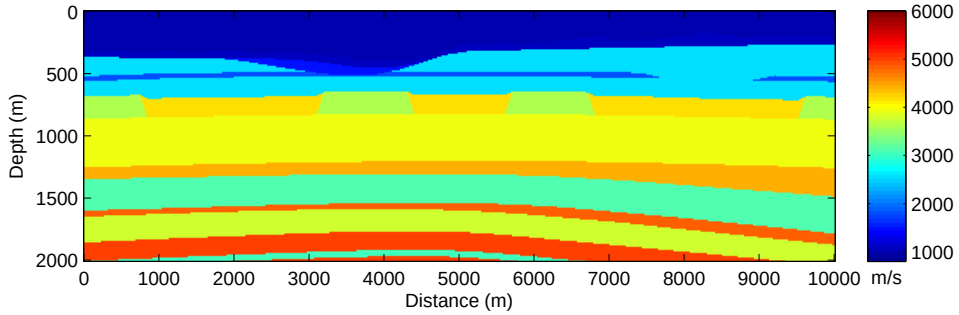
We can see that the estimated result in Figure 4.10(b) matches closely the true result. The near surface is recovered accurately and more details about the curvature of the



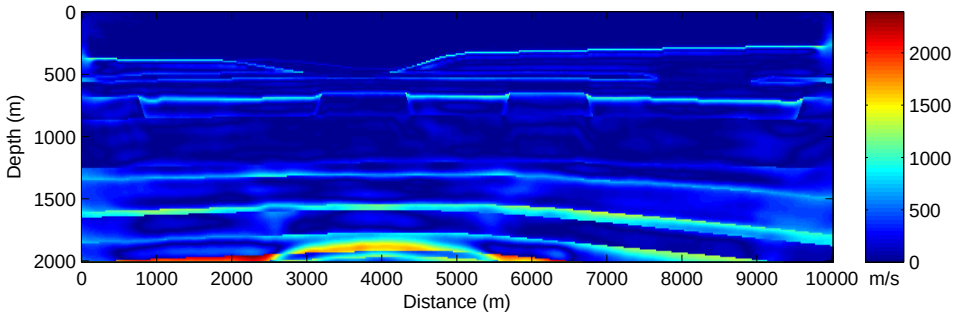
(a) Initial Model



(b) Estimated Model



(c) True Model



(d) Absolute Value of the Difference Between (b) and (c)

Figure 4.6: Preliminary inversion using the sequential strategy. The frequencies inverted are 2-14.8 Hz with 0.2 Hz interval. The initial model is the true model smoothed by 240 m x 240 m smoother.

structure are recovered. Note that the thin low velocity layer, which is a very hard feature to recover, is almost fully recoved, including the truncations. Note also that the feature within the wadi and the low velocity zones within the Rus formation are also recovered.

Figure 4.10(d) shows the absolute value of the difference between the estimated and the true model. Most of regions of the model are recovered within less than +/- 500 m/s. Note that the feature in the bottom of the model is caused by the thin layer at the very edge of the model. Edge effects combined with this thin layer caused this sort of artifact. Understanding such artifacts is crucial in assigning confidence levels to each region of the model and subsequently smoothing or keeping such features.

One-dimensional profiles of the models in three locations are shown in Figure 4.11. You can see that the initial model is an averaged version of the true model. Note how closely the estimated result matches the true result. The estimated models are less accurate in the thin layers. Recovering such small details would require inverting higher frequencies.

We have calculated the one way vertical traveltimes from the acquisition surface to the datum shown in Figure 4.12. Figure 4.12 shows the statics, vertical one-way traveltimes, calculated in the models. Figure 4.15(a) shows the improvement we are able to achieve. The statics in the initial model can be higher than +/- 40 ms. The statics calculated in the estimated model are less than +/- 10 ms. With this much accuracy of estimating the statics, residual statics methods would not have difficultly correcting for the rest of statics.

4.4.2 Efficient Strategy

In this experiment, we use more efficient strategy of (Sirgue and Pratt, 2004a). The efficient strategy suggests inverting only four frequencies: 2, 4.5, 10, 22.4 Hz. We used

30 iterations per frequency. The total number of iterations is 120. Figure 4.16 shows the frequencies selected and their relation with the wavenumber. The selection strategy is based on selecting those frequencies whose wavenumber range in total would span the wavenumber domain. The redundancy of the data is fully utilized through the efficient strategy.

Figure 4.13 shows the results. Note that we are able to recover the low velocity zone, the wadi's low velocity zone and in a bit less detail the Rus collapses. Note also that low relief structure especially at Wasia formation is recovered in more details.

Although the result here seems to be less accurate, a closer look at the difference plot shows that the resulting model is almost within +/- 500 m/s in many areas. A closer look at the 1D profiles in Figure 6 shows how close the match is. The resolution of the recovered model is comparable to that of the sequential approach for almost 37.5% of the cost.

Obtaining more accurate models using the efficient strategy can be accomplished. However, it would require more human input and interpretation through the application of a more aggressive preconditioning scheme. A big advantage of the efficient strategy is that it can be used as a tool for quick testing of an array of parameters. For such a computationally intensive process, the efficient strategy can make a big difference.

We have calculated the statics for the models as shown in Figure 4.15. Note that the statics are less accurate than the previous experiment. Nevertheless, in most regions they are less than +/- 15 ms, which is accurate enough for residual statics to be effective when applied afterward.

Figure 4.5 shows three shot records in different locations: left, middle and right-hand side of the model. The shots generated using the true model (observed data), the estimated and the difference between them. Note how close the match is between the observed and the predicted data. The difference between them seems to be a scaled

version of the observed data. This kind of match is what one should look for when using real data.

4.5 Inversion from a higher frequency

Now, we will take a look at the results starting from a higher frequency. We use identical parameters as we have used previously but the only difference is that we start the inversion from a higher frequency, 5 Hz.

4.5.1 Sequential strategy

Using the sequential strategy and inverting the frequencies starting from 5 till 14.8 Hz with 0.2 Hz interval, we get the result shown in Figure 4.17(b) and the associated 1D profiles in Figure 4.18. The initial model is the same that used in the final inversion, namely the true model smoothed by a 480 m x 480 m smoother.

In order to better compare this result with previous ones, the color bar of the estimated result is scaled up to 6000 m/s although the highest result could reach about 10000 m/s, as can be seen in Figure 4.17(d). We can note that the result is not as accurate as the previous ones. As a matter of fact, small features, which are artifacts, are very visible in the resulting model. Those seem to be indicative of possible divergence from the global minimum. What is even more interesting is the fact that the thin layer above about 700 m to the left of the model is very visible. But quantitatively the layer is higher in velocity rather than its soundings. A particular area of even more divergence is the wadi area. The velocities there are more overestimated than any other area as shown in Figure 4.17(d).

As mentioned previously, perhaps more accurate results can be obtained though more aggressive preconditioning of the updates or the gradient of the misfit function. But the issue is that much more interpretive analysis will need to be imposed on the result. In

other words, the result will be more subjective.

The result in Figure 4.17(b) shows some indications of the artifacts, as well as, the possible divergence from the global minimum. Small features show up in the result including smile-shaped ones, which are typical of migration-based methods. In addition to the small features, we were able to detect a thin layer in some areas but the quantitative result is not accurate. The Wadi area as the least to be quantified accurately as shown also in Figures 4.18(a) and 4.17(d). More attention should be paid to the Wadi area since more artifacts result from it, possibly because of its relatively low elevation compared with the surrounding areas.

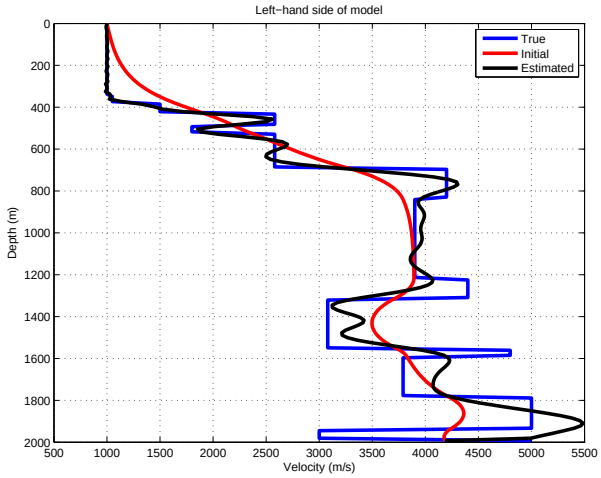
4.5.2 Efficient strategy

For the efficient strategy the same analysis applies as in the previous section. But what is of particular interest is that the result of using the efficient strategy is close to that of the sequential strategy even in the case of possible divergence as shown in Figures 4.19(b) and 4.20. Although the resulting model has more small features, the thin layer on top is very close to that of the sequential strategy.

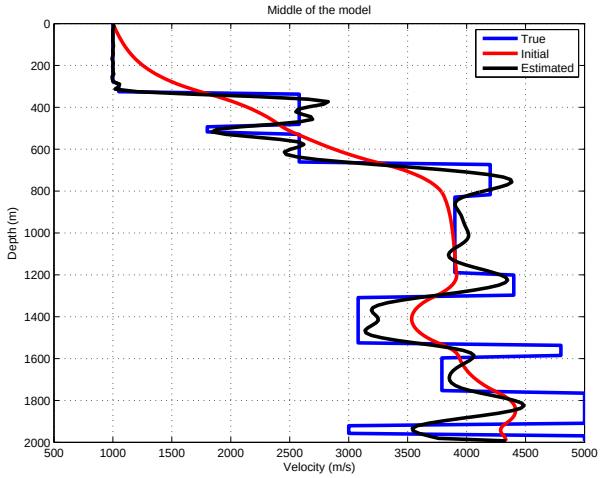
4.6 Conclusions and Discussion

We have applied waveform inversion on a realistic model of an area of complex near surface and low relief structure. Frequency domain waveform inversion successfully recovered the compressional velocity of the model using a smooth initial model. Waveform inversion was able to recover the true model in using a sequential approach and less accurately using the efficient strategy. The statics solutions in both cases are accurate enough for residual statics to be effective. We have observed that in the case of starting from a higher frequency, artifacts of small features are often visible especially when using the efficient strategy. The result of the efficient strategy in this case even more closely

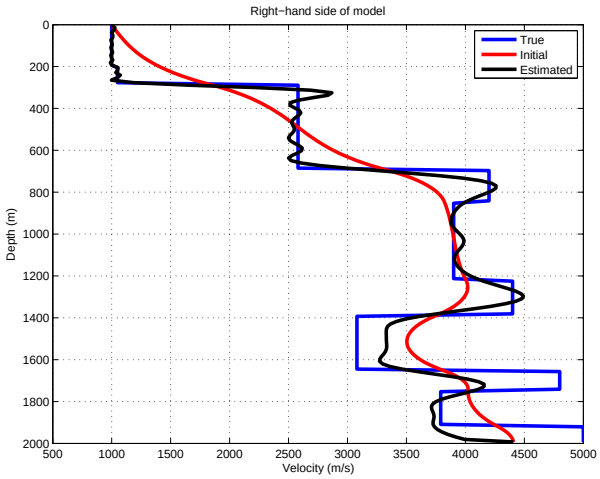
resembles that of the sequential strategy.



(a)

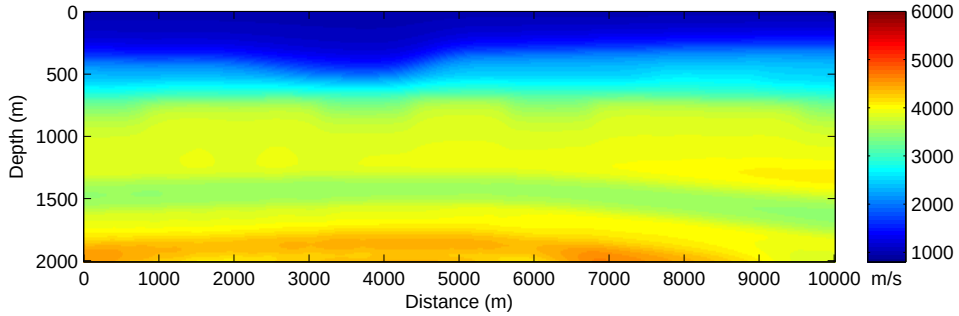


(b)

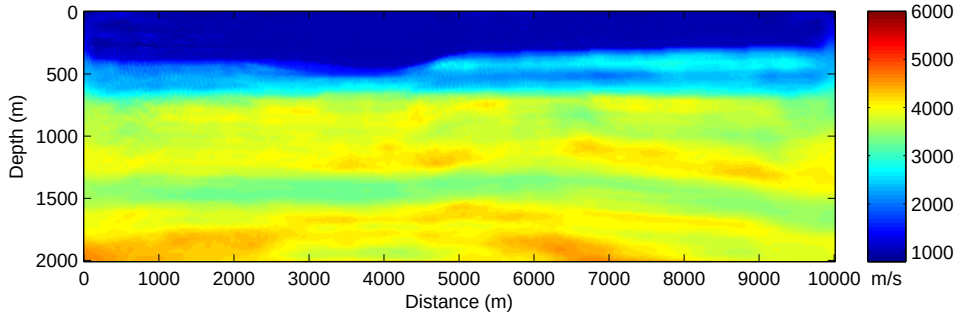


(c)

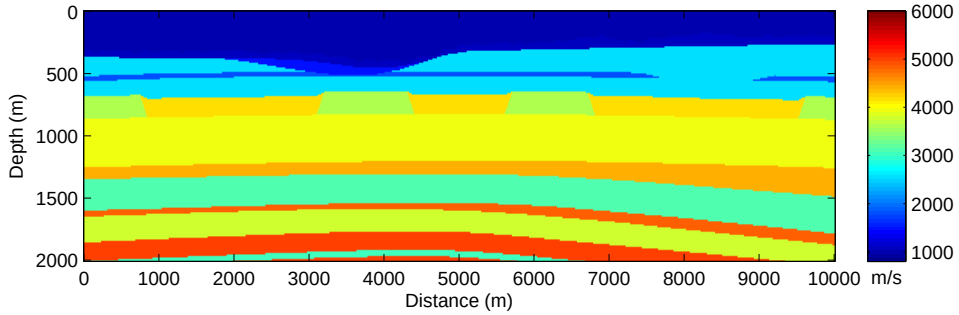
Figure 4.7: 1D velocity profiles from the preliminary inversion using the sequential strategy. The profiles are from 3 different locations.



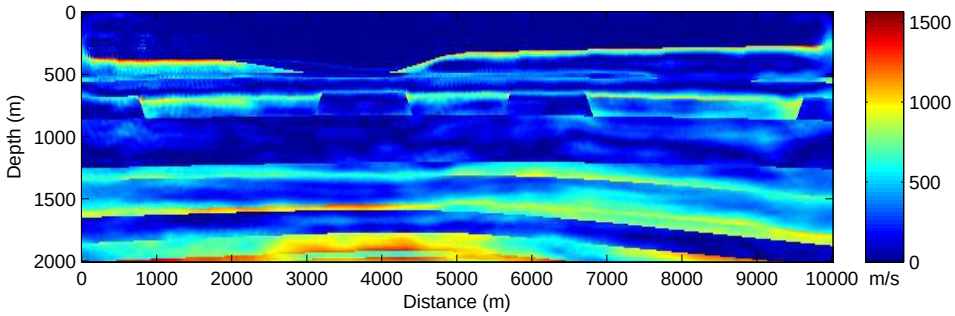
(a) Initial Model



(b) Estimated Model

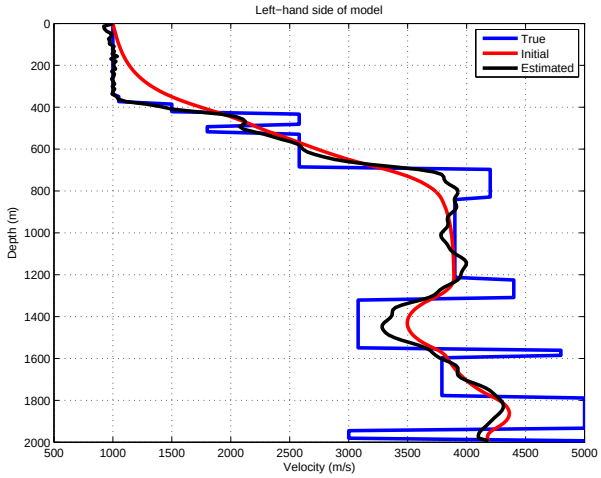


(c) True Model

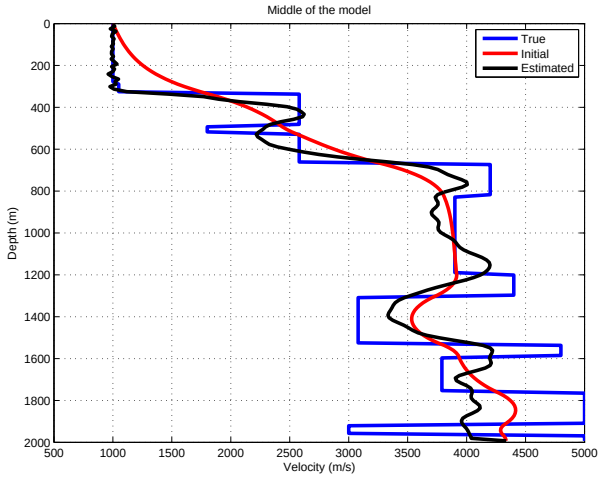


(d) Absolute Value of the Difference Between (b) and (c)

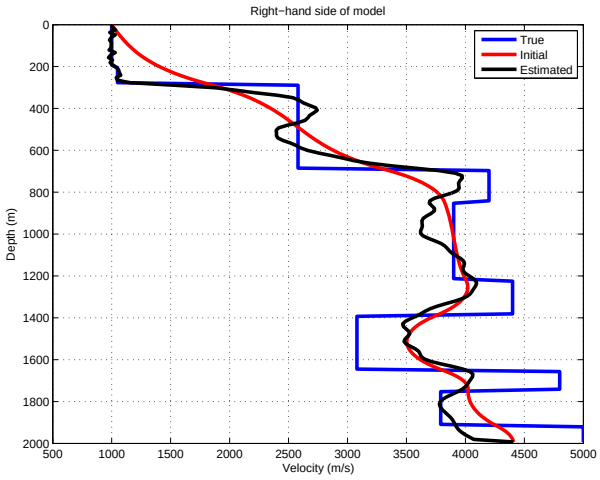
Figure 4.8: Preliminary inversion using the efficient strategy. The frequencies inverted are 2, 4.5, 10 and 22.4 Hz. The initial model is the true model smoothed by 240 m x 240 m smoother.



(a)

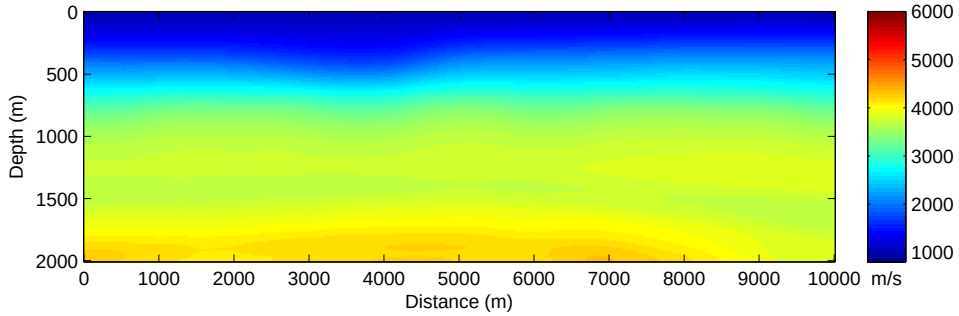


(b)

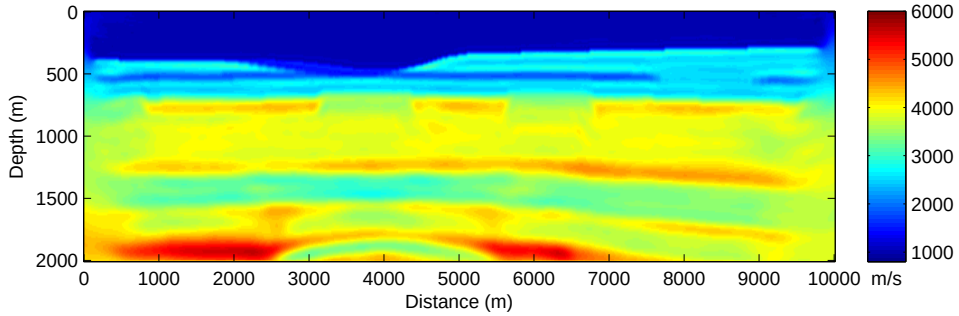


(c)

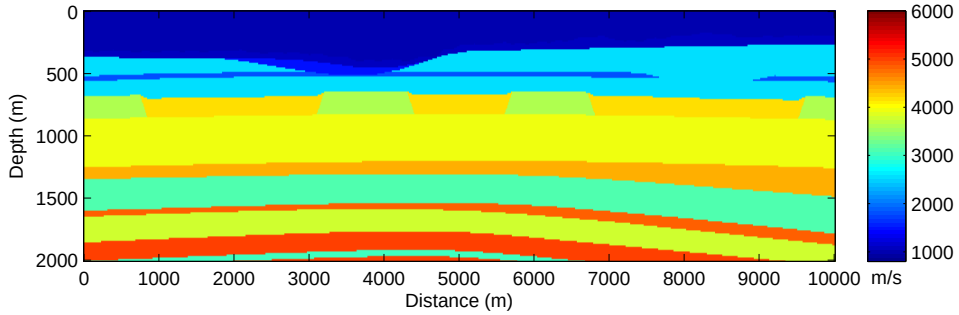
Figure 4.9: 1D velocity profiles from the preliminary inversion using the efficient strategy. The profiles are from 3 different locations.



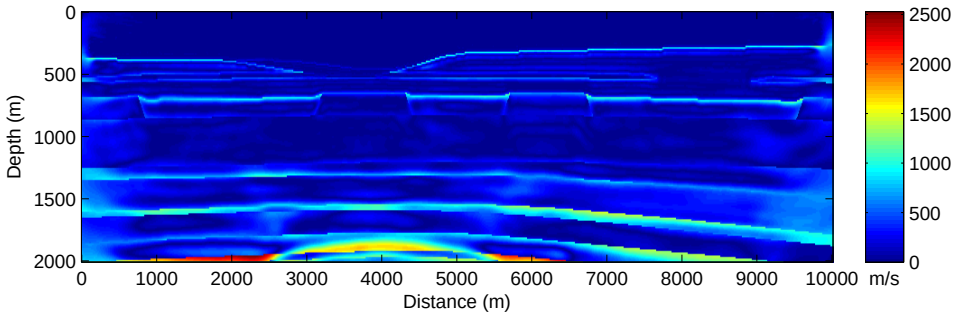
(a) Initial Model



(b) Estimated Model

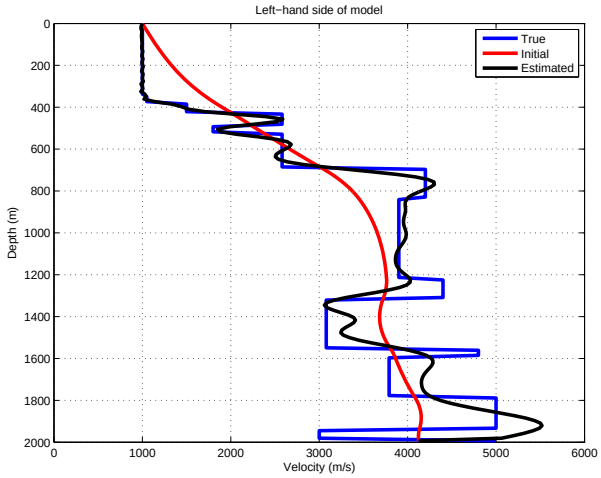


(c) True Model

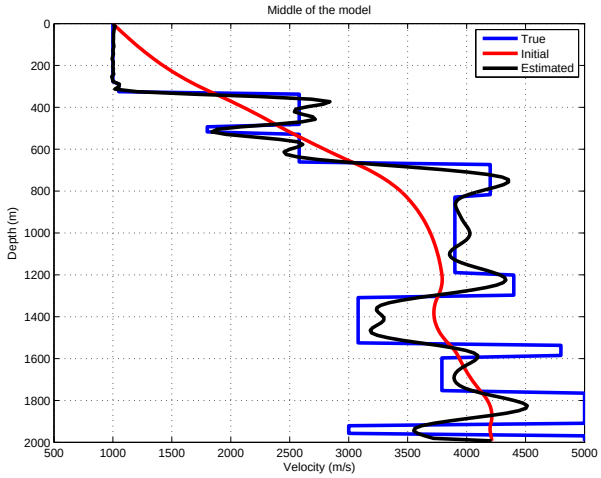


(d) Absolute Value of the Difference Between (b) and (c)

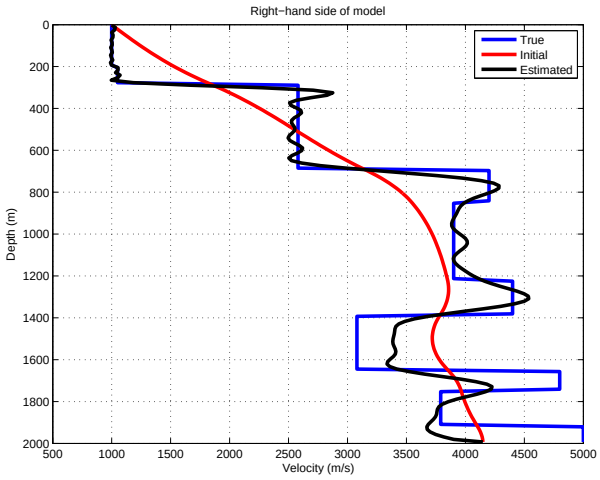
Figure 4.10: Final inversion using the sequential strategy. The frequencies inverted are 2-14.8 Hz with 0.2 Hz interval. The initial model is the true model smoothed by 480 m x 480 m smoother.



(a)

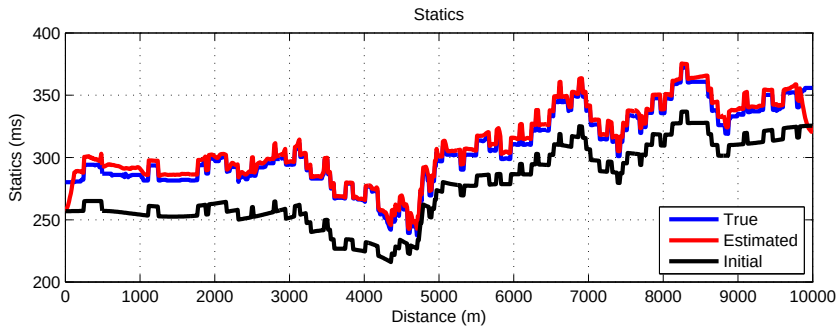


(b)

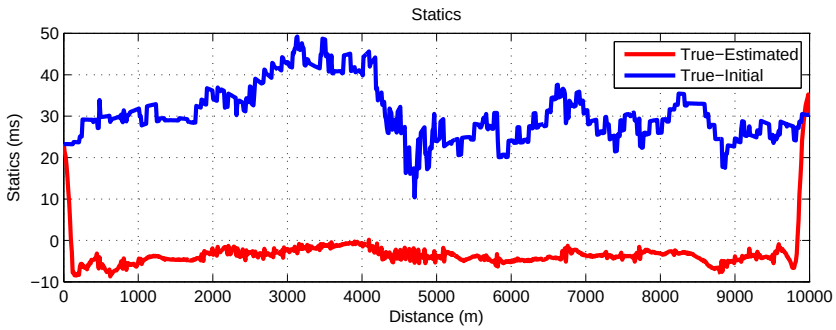


(c)

Figure 4.11: 1D velocity profiles from the final inversion using the sequential strategy. The profiles are from 3 different locations.

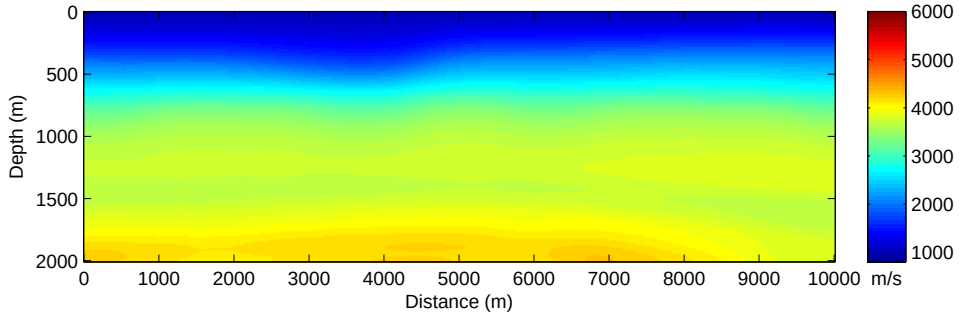


(a)

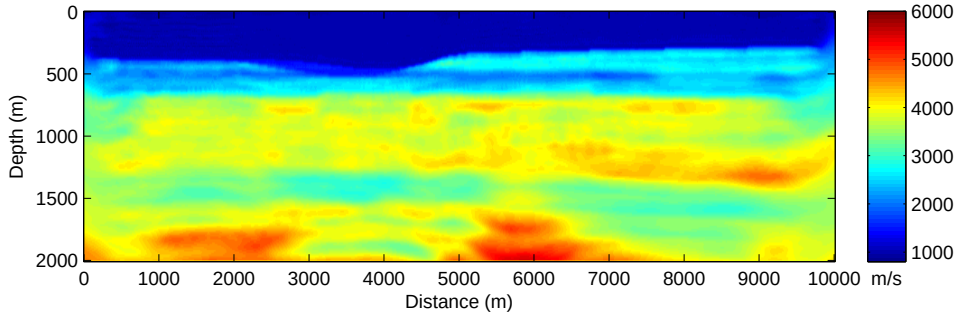


(b)

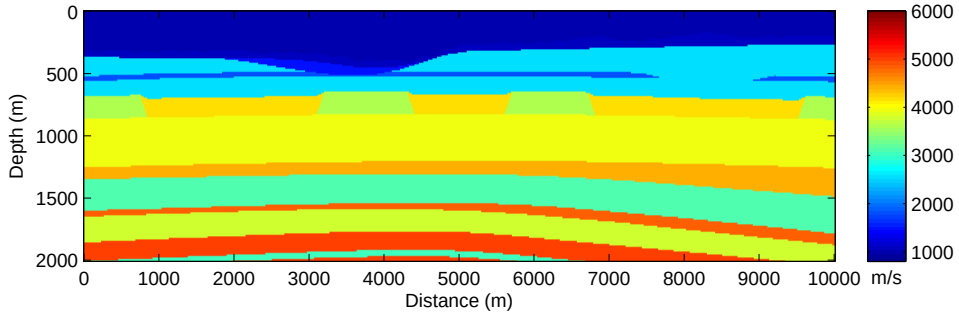
Figure 4.12: Statics for the sequential experiment. a) shows all the statics b) shows the differences.



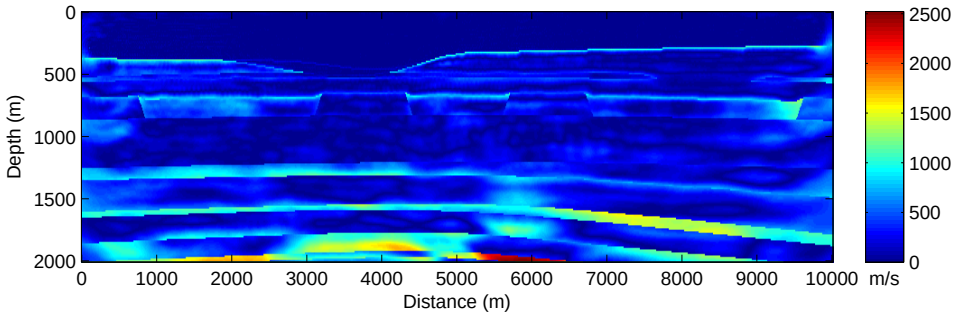
(a) Initial Model



(b) Estimated Model

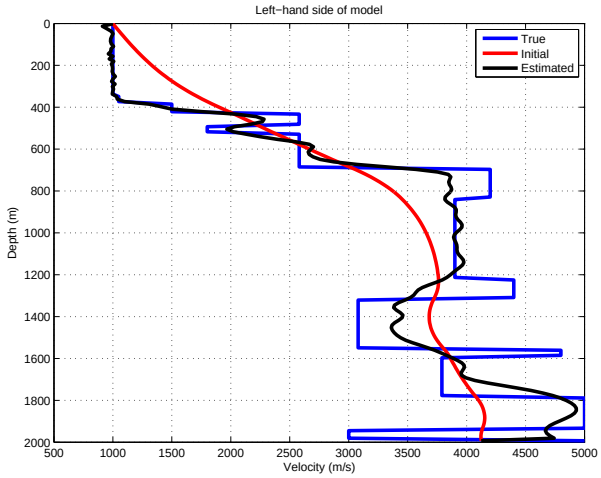


(c) True Model

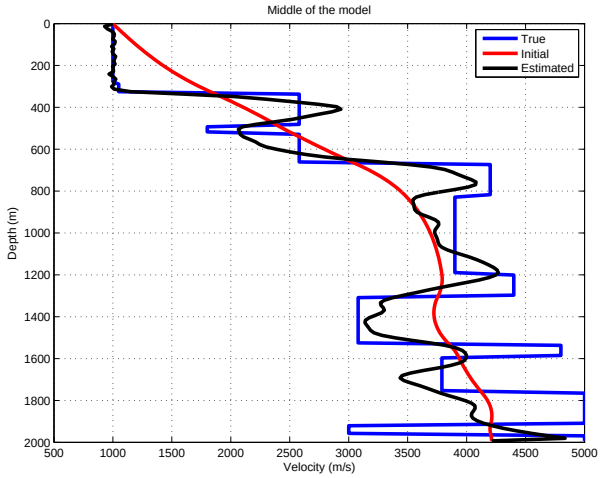


(d) Absolute Value of the Difference Between (b) and (c)

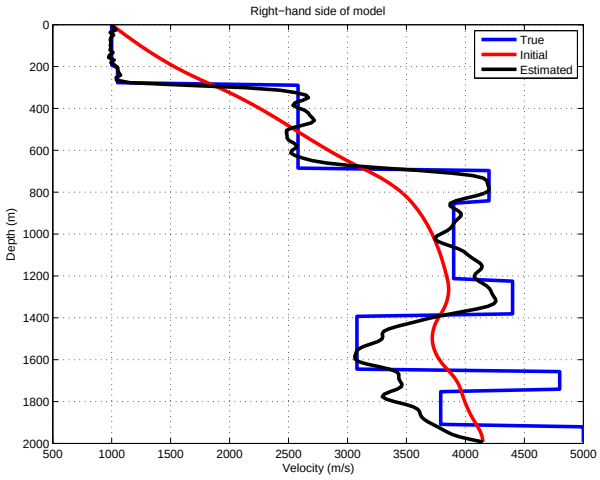
Figure 4.13: Final inversion using the efficient strategy. The frequencies inverted are 2, 4.5, 10 and 22.4 Hz. The initial model is the true model smoothed by 480 m x 480 m smoother.



(a)

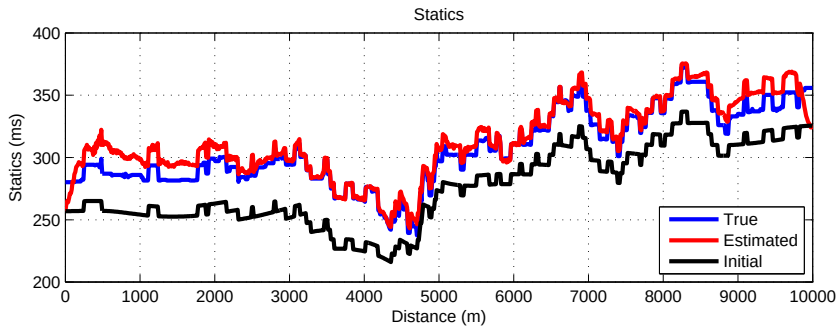


(b)

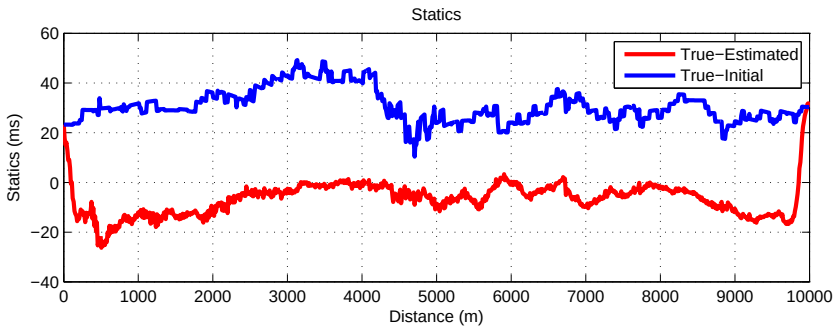


(c)

Figure 4.14: 1D velocity profiles from the final inversion using the efficient strategy. The profiles are from 3 different locations.



(a)



(b)

Figure 4.15: Statics for the efficient experiment. a) shows all the statics b) shows the differences.

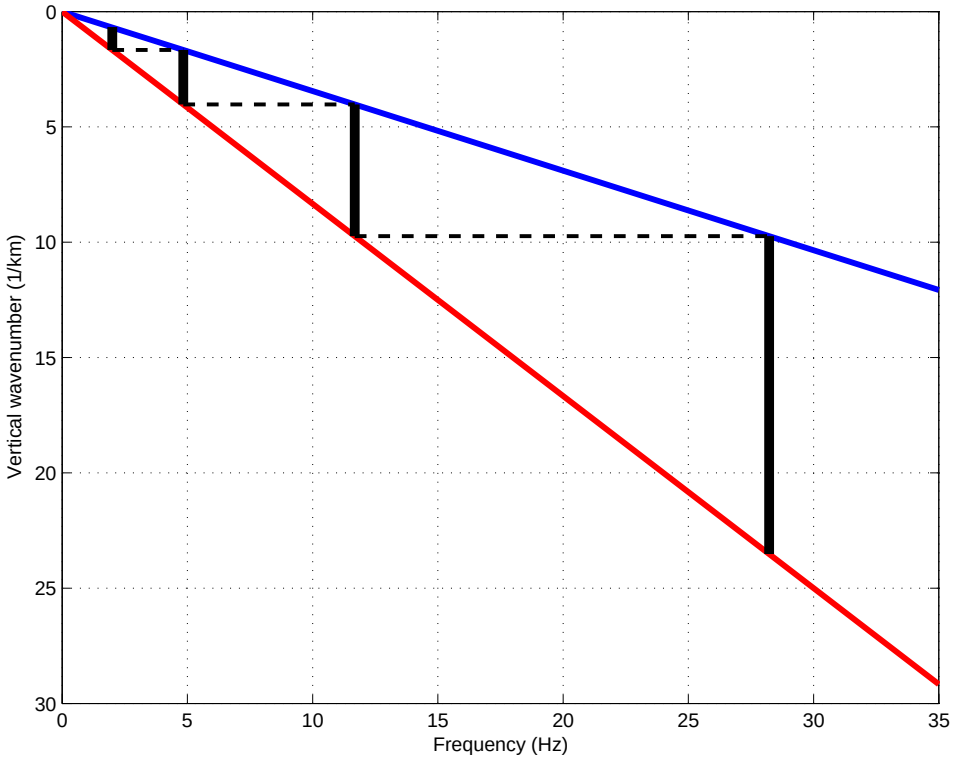
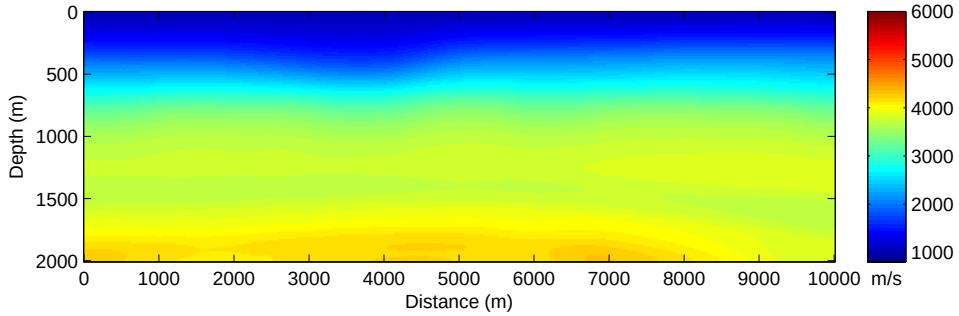
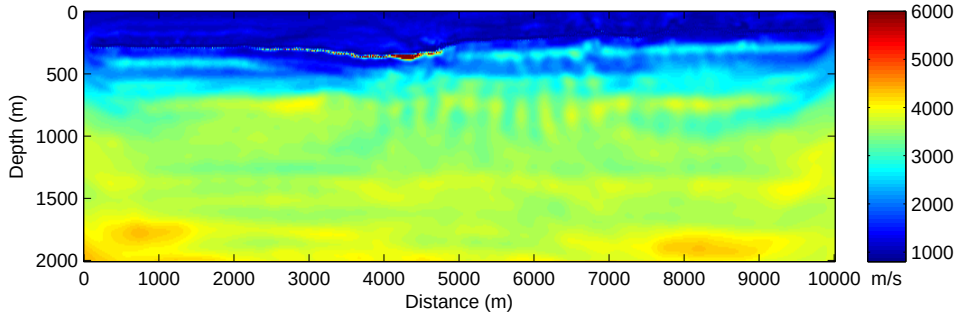


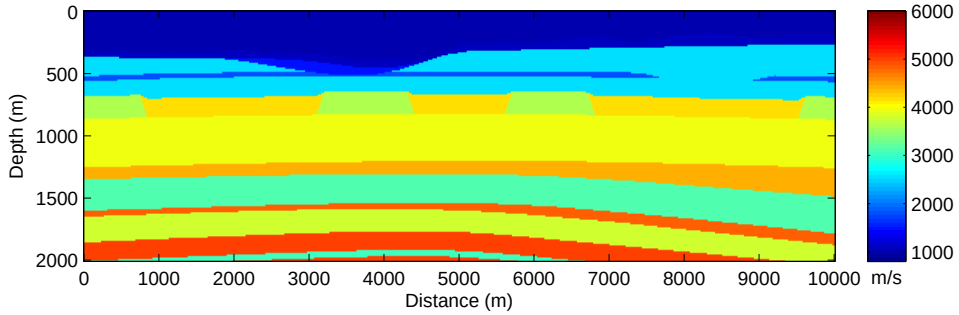
Figure 4.16: The frequencies selected for the efficient strategy.



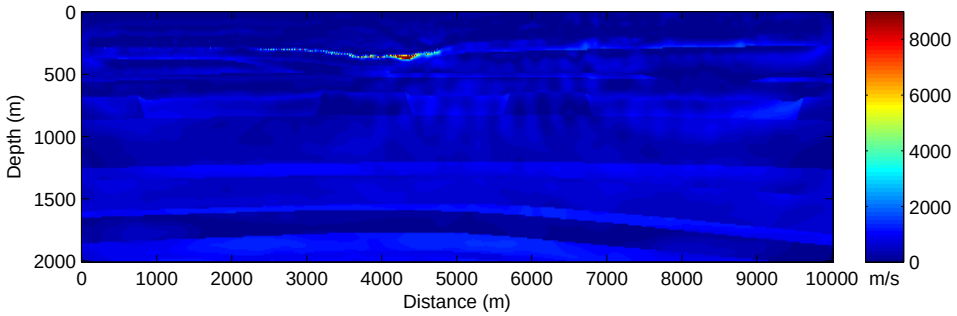
(a) Initial Model



(b) Estimated Model

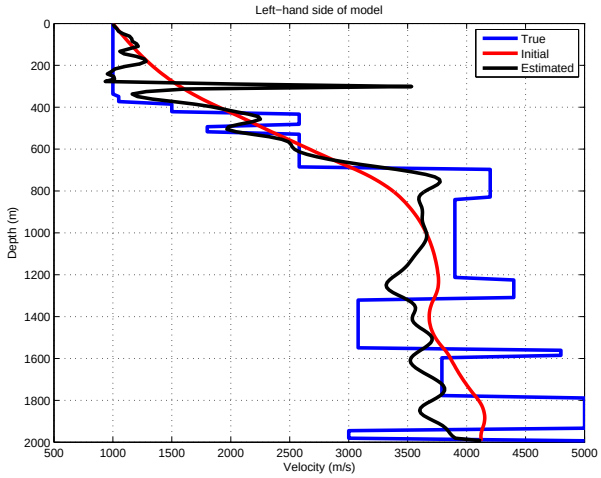


(c) True Model

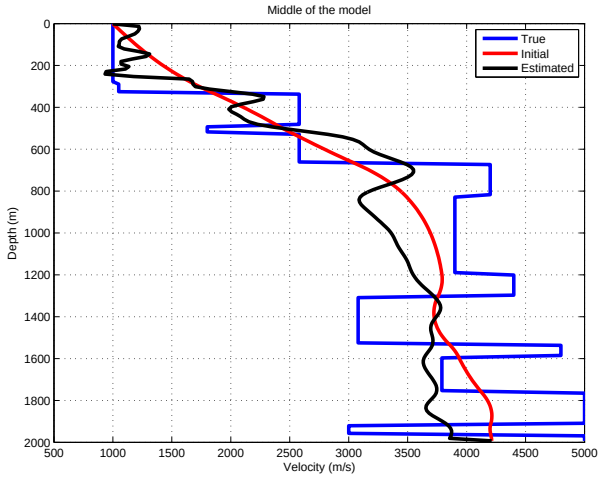


(d) Absolute Value of the Difference Between (b) and (c)

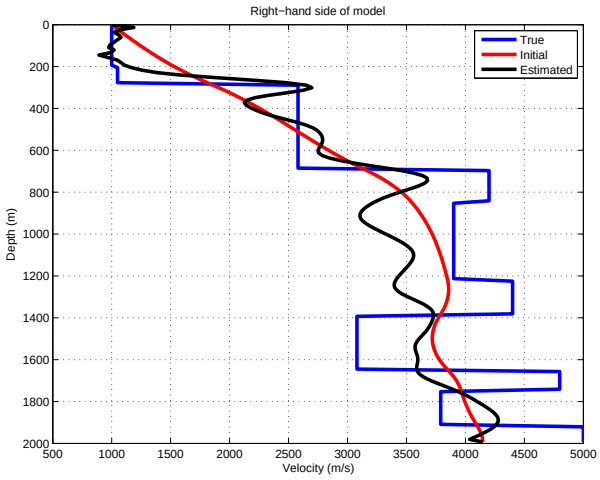
Figure 4.17: The models associated with the inversion starting from a higher frequency and using the sequential strategy. The frequencies inverted are 4-14.8 Hz. The initial model is the true model smoothed by 480 m x 480 m smoother.



(a)

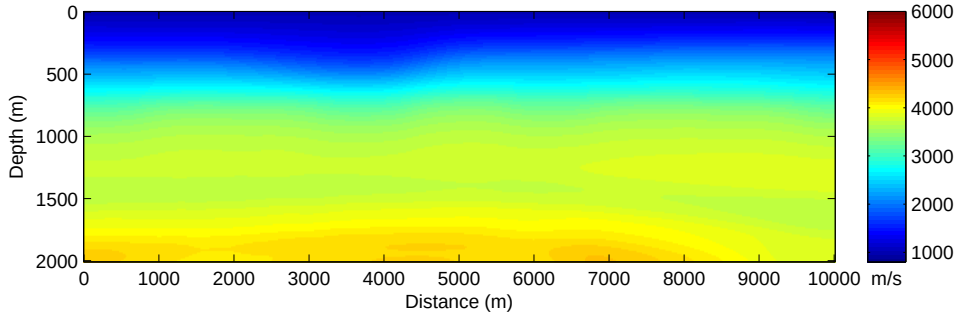


(b)

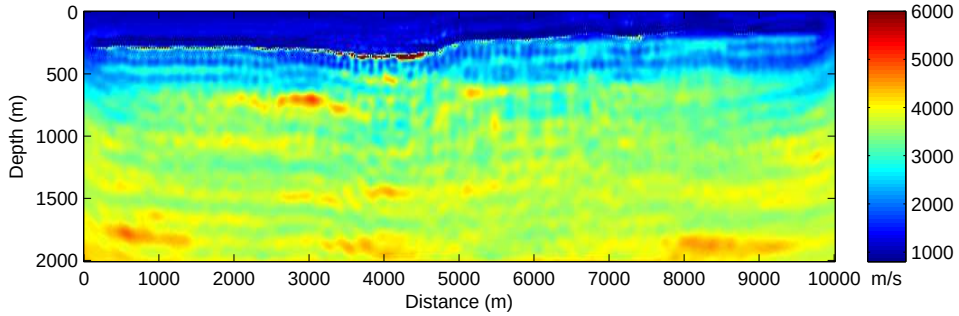


(c)

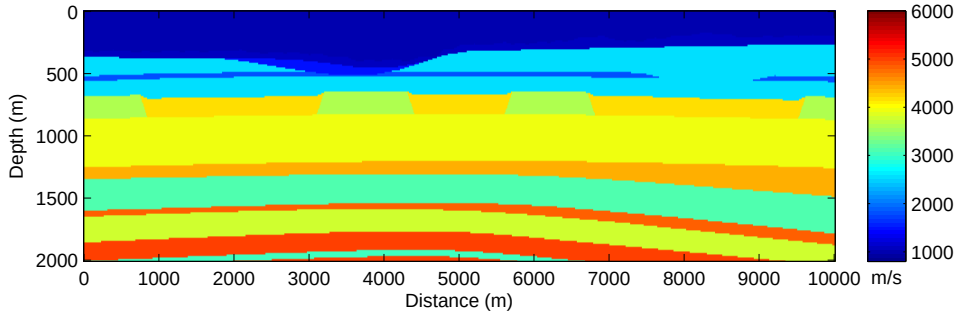
Figure 4.18: 1D velocity profiles from the inversion starting from a higher frequency and using the sequential strategy. The profiles are from 3 different locations.



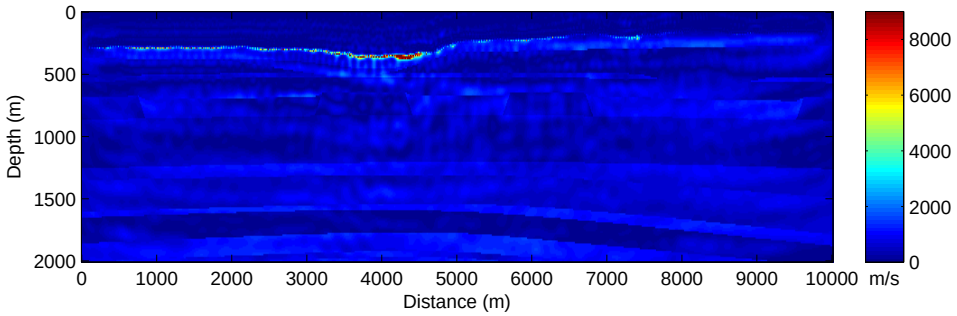
(a) Initial Model



(b) Estimated Model

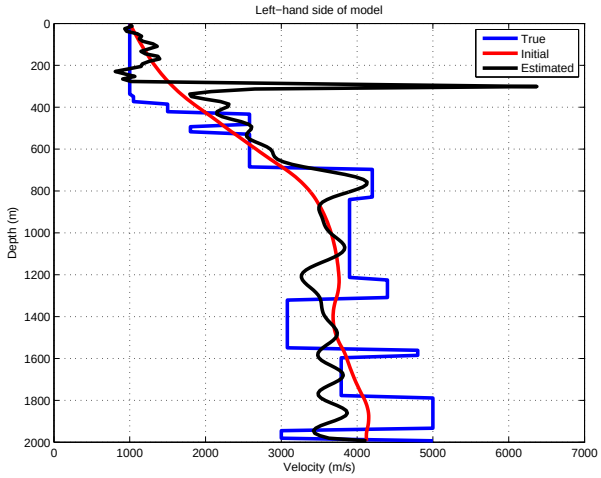


(c) True Model

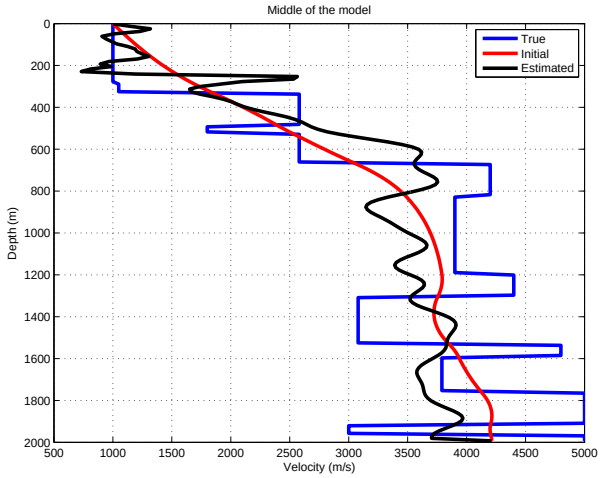


(d) Absolute Value of the Difference Between (b) and (c)

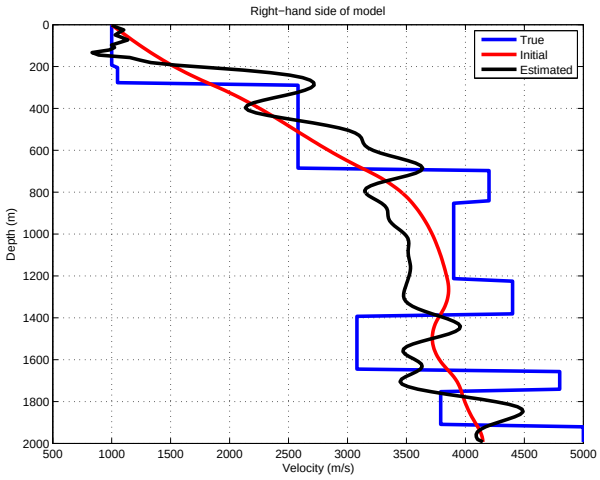
Figure 4.19: The models associated with the inversion starting from a higher frequency and using the efficient strategy. The frequencies inverted are 5 and 11.2 Hz. The initial model is the true model smoothed by 480 m x 480 m smoother.



(a)



(b)



(c)

Figure 4.20: 1D velocity profiles from the inversion starting from a higher frequency and using the efficient strategy. The profiles are from 3 different locations.

This page is intentionally left blank.

Chapter 5

Waveform inversion and shingling

Seismic data possessing first arrival shingling is indeed difficult to interpret and infer the velocities of their models. In this chapter, we attempt to explain the phenomenon of shingling. As far as we know, there is currently no known general solution to this problem. We will address the question of whether waveform inversion can indeed resolve the velocity structure responsible for shingling.

5.1 Introduction

First arrival shingling is the phenomenon associated with the first arrivals possessing an en echelon or stair-step shape. Cox (1999) defines it in his book of near surface geophysics, as "a phenomenon characterized by a shift of energy to successively later cycles as the offset increases producing an en echelon pattern" (Cox 1999, p. 158). In fact, Sheriff and Geldart (1995) in their textbook also define and discuss this phenomenon. They define it as "the shift of peak energy to later in the wavetrain, as the offset distance increases and the number of cycles increases" (Sheriff and Geldart, 1995, p. 169).

So it is a rather unique way for the amplitudes to decay with offset producing an en echelon pattern. The term 'shingling' seems to first appear in the literature in the paper by Press et al. (1954); however, they attribute the naming to another person. Another term that is used interchangeably but perhaps less commonly is *echeloning* as used by Press et al. (1954), for instance.

Figure 5.1 shows an example from Cox (1999). Note that if one were to pick a curve of the first arrivals, the curve would look blocky and shaped like a stair step .

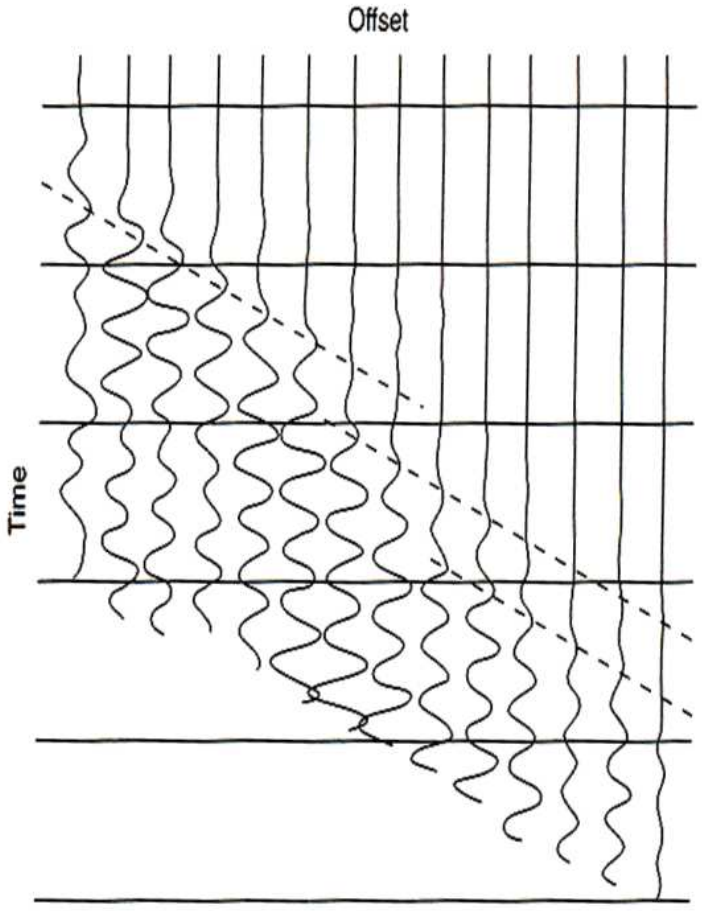


Figure 5.1: An example of first arrival shingling from Cox (1999).

First arrival shingling is a very long lasting issue. It is known for many decades now. First arrival shingling makes it very difficult to interpret the first arrivals accurately enough and therefore to infer the Earth's velocities.

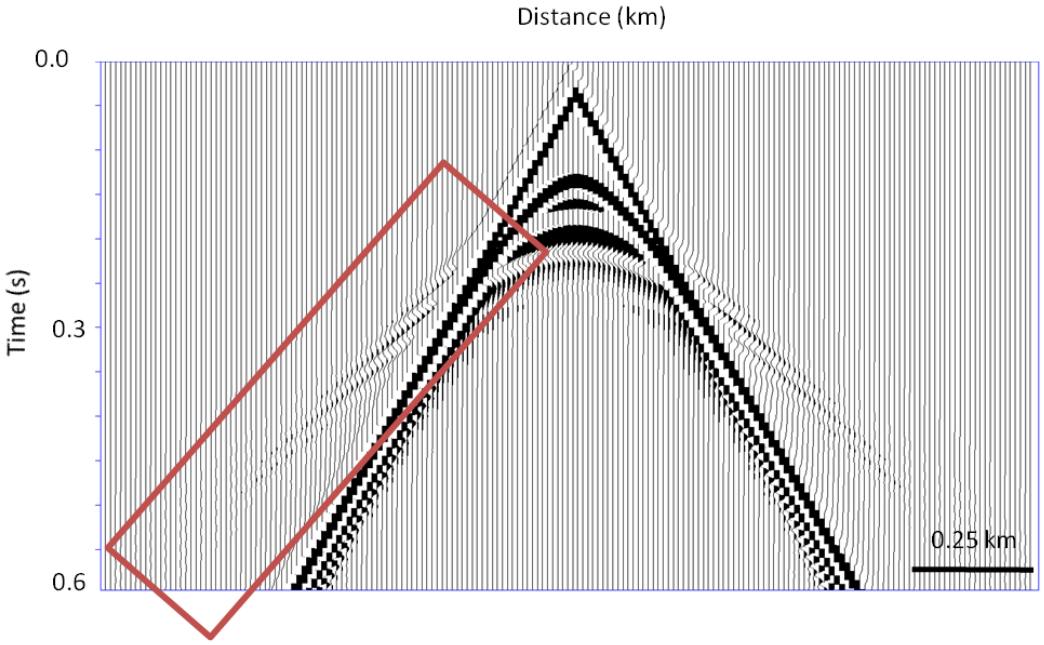
Before seismic reflection surveys were common, seismic refraction surveys were more common. The refraction surveys rely heavily on refractions, mainly head waves and/or diving waves, as the name implies. Therefore, such a problem as shingling posed a serious obstacle for reliable interpretation.

Some studies have been conducted to further understand this phenomenon. Most of those studies relied on physical modeling experiments, possibly due to the expense of numerical modeling at the time. Press et al. (1954) have shown that the refracted arrivals decay rapidly if the beds are thin. They showed this using a physical modeling experiment that incorporated such materials as Plexiglass , brass and aluminum.

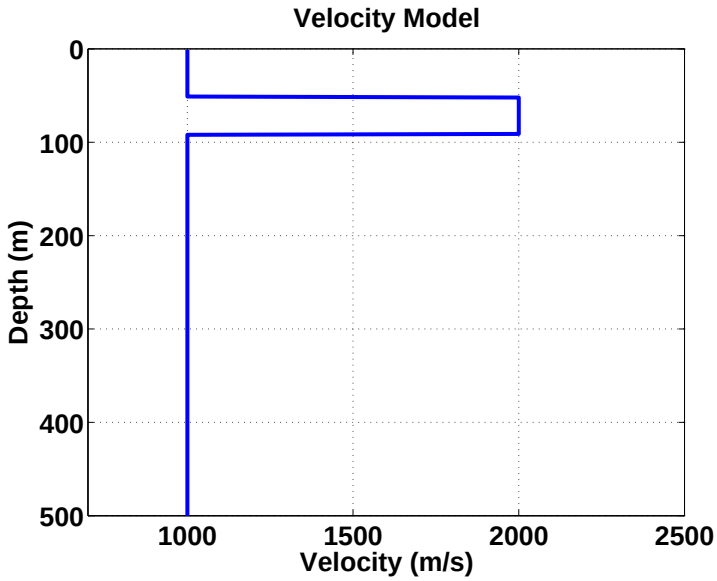
Lavergne (1961) conducted another physical modeling experiment and showed the difficulties associated with quantifying the velocities of thin layers. Lavergne (1966) attempted to quantify the thickness of the beds. They showed that if the refactor is thinner than $\lambda/6$ significant decay would result. Poley and Nooteboom (1966) studied thin fast layers which significantly reduce the penetration of energy to further depths. *Screening effect* is the term they used to describe such a phenomenon. Given the signal to noise ratio at the time, this posed as a serious issue.

Rosenbaum (1965) conducted a physical modeling experiment, as well. He showed that the shingling effect can indeed result from thin beds. Spencer (1965) has also shown that the interference of multiples with the the head wave and primary reflection can indeed cause shingling. Figure 5.2(a) shows a numerical example illustrating such interference resulting from a model that contains a thin fast layer as shown in Figure 5.2(b). We used a Ricker wavelet to generate the shot with 30 Hz dominant frequency.

Although those studies have explain some aspects of the this phenomenon, they have



(a) Shot record



(b) Velocity model (1D profile)

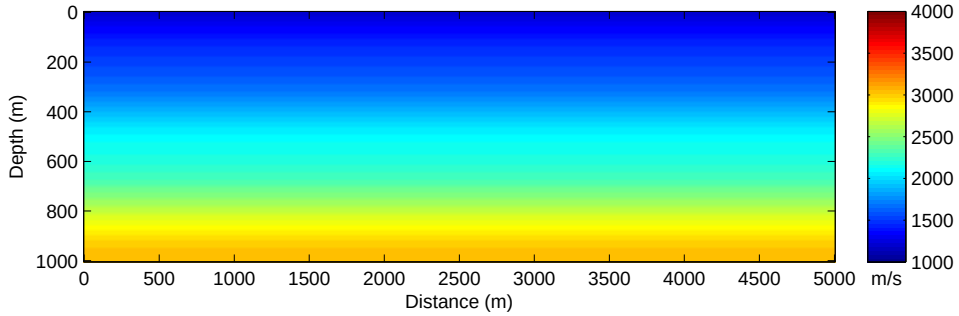
Figure 5.2: An example of first arrival shingling in the rectangle in (a). The shingling is caused by a thin fast layer shown in (b). The reflections including the multiple reflections interfere with the head wave producing the shingling effect.

not resulted in any mainstream technique to tackle this issue, as far as we are aware. Since the kinematics of shingling get more obscured and complex, they don't always seem to be sufficient to come up with a consistent answer. Spencer (1965) studied the amplitude decay theoretically and proposed an analytic modeling formula using geometric ray theory.

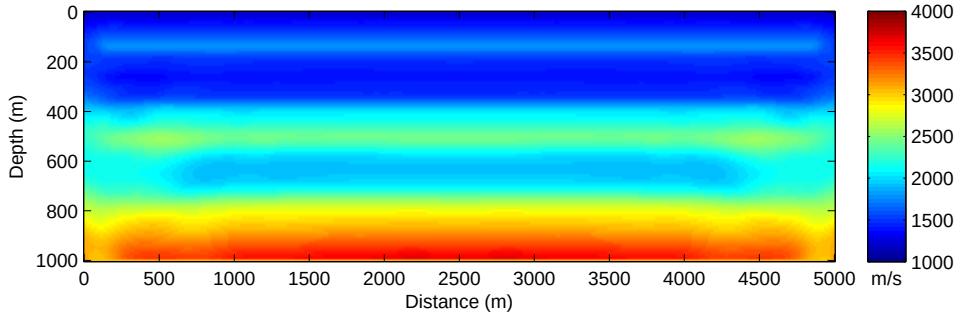
Therefore, those studies have shown that thin beds cause the shingling effect since they cause the refracted arrivals to decay much more rapidly and may cause the reflections, including multiple reflections, to interfere with the refractions, as well.

Another scenario that causes shingling is shown in Figures 5.3(c) and 5.4 . The model shows two low velocity zones sandwiched between layers whose velocities increase linearly with depth. The model consists of a top layer of velocity increasing linearly with depth. Below that is a low velocity zone. Beneath that is a layer whose velocity increases with depth, followed by a low velocity zone and then another layer whose velocity is increasing linearly with depth.

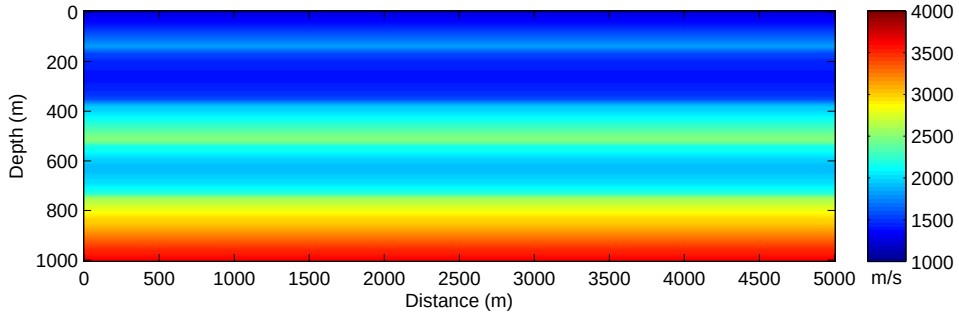
A shot record was created for this model (with a ricker wavelet of 30 Hz dominant frequency) and is shown in Figure 5.5. The first arrivals first decay and then appear again and decay once again. To understand this further, we traced rays through the model as shown in Figure 5.6. Notice that there exists some shadow zones where no diving rays reach the surface. The presence of such shadow zones caused the shingling of the first arrivals.



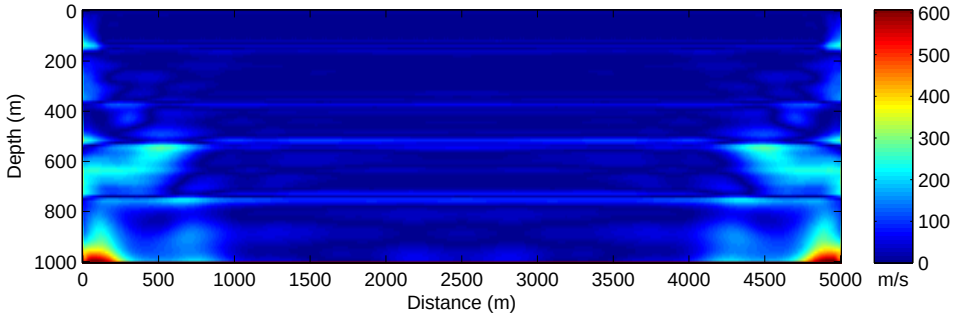
(a) Initial model



(b) Estimated model

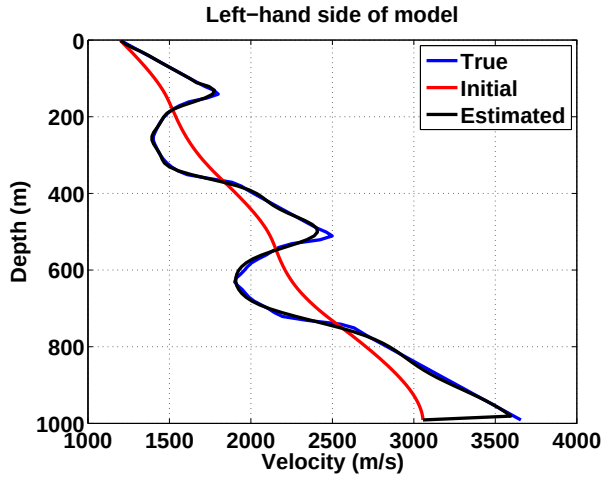


(c) True model

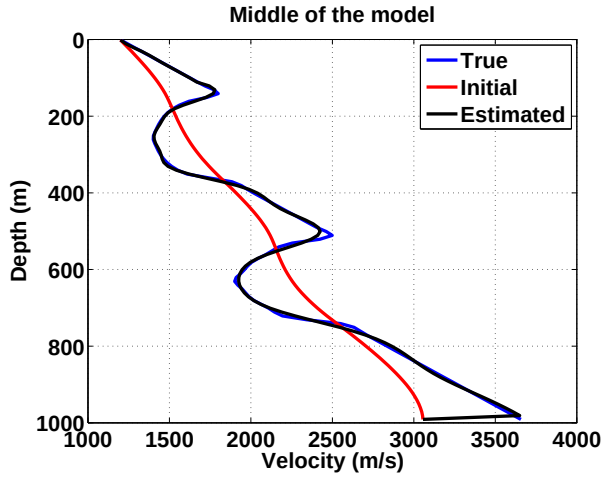


(d) Absolute value of the difference between a) and c).

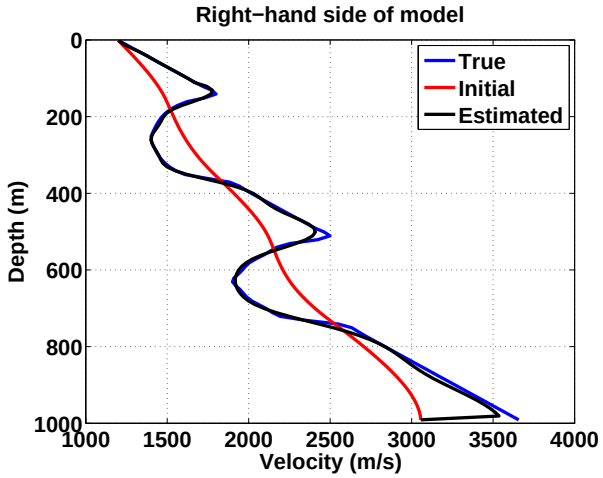
Figure 5.3: The true model that produces shingled first arrivals the its associated models resulting and used in waveform inversion.



(a)



(b)



(c)

Figure 5.4: 1D velocity profiles for the models shown in Figure 5.3.

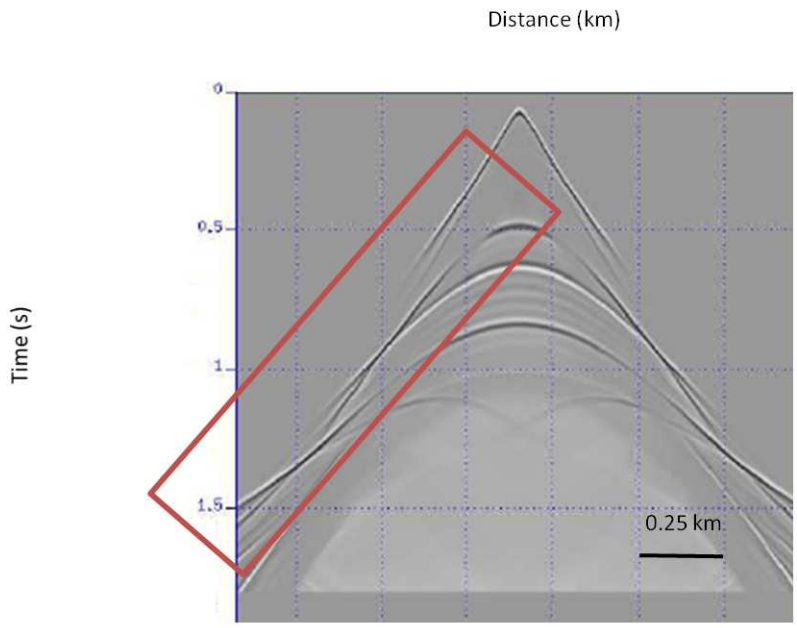


Figure 5.5: An example of first arrival shingling (in the rectangle) caused by two separate low velocity zones sandwiched between layers whose velocities increase linearilly with depth as shown in Figure 5.4.

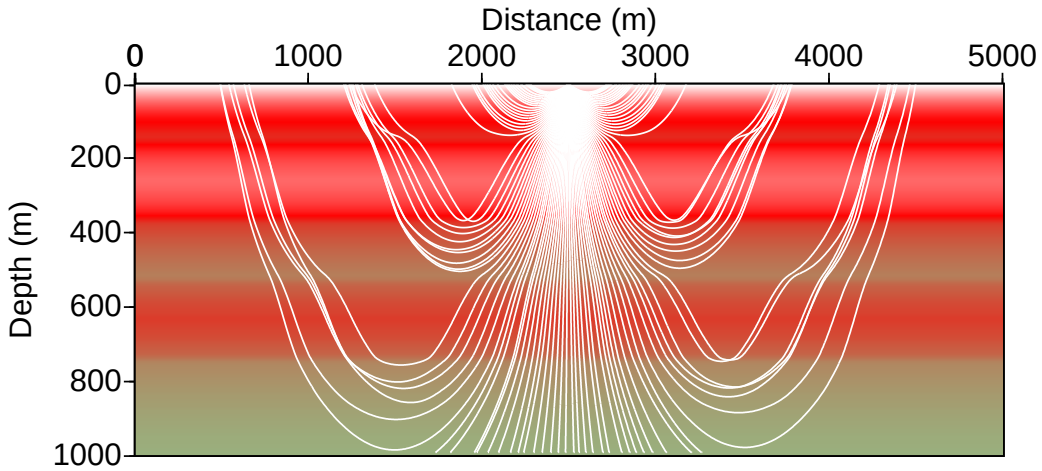


Figure 5.6: Rays traced in the true model shown in Figures 5.3(c) and 5.4. Notice the shadow zones, where no diving rays reach the surface.

5.2 Waveform Inversion

We have applied waveform inversion to invert for the shingling model shown in Figures 5.3(c) and 5.4. We first created the synthetic data of 191 shots and each has 240 receivers with 25 m shot spacing and 20 m receiver spacing. The first shot is placed at 115 m and the first receiver is placed at 110 m from the left-hand side of the model. The initial model used is the true model smoothed by 300 m x 300 m smoother. The frequencies inverted are 2-15Hz with 0.2 Hz interval and the gradient of the misfit is preconditioned by filtering the wavenumber domain. The results are shown in Figures 5.3(b) and 5.4. The result illustrate that waveform inversion does not struggle at all in successfully inverting for the velocity.

5.3 Conclusion

We have attempted to explain and discuss the phenomenon of shingling in this chapter. We applied waveform inversion to a realistic model that causes the shingled first arrivals. We have shown that waveform inversion can indeed resolve such models.

This page is intentionally left blank.

Chapter 6

Conclusions and discussion

We started the thesis with chapter one, where we have given an overview of the thesis and the fields of seismic migration and inversion. In chapter two, we discussed the theoretical aspects of waveform inversion. We emphasized the fact that waveform inversion is based on the well-known linearization, the Born approximation, and that in turn makes it based on the well established method of seismic migration. Waveform inversion is similar to traveltime tomography with one main difference: traveltime tomography uses the high frequency thin raypaths whereas waveform inversion uses the finite-frequency band-limited wavepaths.

Waveform inversion for homogeneous media is roughly the same as Kirchhoff migration and kinematically equivalent. True amplitude inversion of Bleistein et al. (2001) is also similar to waveform inversion with some differences; roughly the weighting of the data is different and the fact that true amplitude inversion uses an approximate Green's functions, namely those based on the WKBJ approximation, whereas waveform inversion typically uses the more accurate numerical methods such as the finite-difference method. Needless to say, waveform inversion is iterative whereas true amplitude inversion is not.

In chapter three, we have shown that strong velocity heterogeneity, a character of complex near surface environments, does not seem to cause issues with waveform inversion. In fact we have inverted the two models, one with strong velocity variations and one without, using the same parameters and both results where highly accurate.

In chapter four, we have inverted a realistic model from a region with highly complex near surface environment. Using realistic preconditioning of the gradient of the misfit function, waveform inversion was able to achieve accurate results starting from 2 Hz.

Starting the inversion from a higher frequency with the same initial model caused waveform inversion to diverge. This implies that in real experiments where low frequencies are to be acquired, it is much safer to use 2 Hz geophones rather than, for instance, their 4.5 Hz counterparts. Although the model we used is indicative of such complex near surface environments, it might be more difficult to invert that even real models due to its high degree of blockiness, a characteristic of models that often causes migration and inversion methods to struggle as noted by Gray et al. (2001).

In chapter five, we have discussed the phenomena of first arrival shingling, a difficult near surface problem. We have demonstrated that waveform inversion can indeed resolve models causing such a phenomenon.

Applying waveform inversion to real data, especially surface seismic data, is indeed a difficult task. This explains the relatively few case studies in the literature, although there are some (e.g Crase et al., 1990; Operto et al., 2004; Sheng et al., 2006). As we mentioned in chapter 1, the main reason why it is difficult is due to the fact that real data often do not contain the necessary information to make waveform inversion work, namely the low frequencies and the large offsets. In addition, other phenomena like attenuation, elasticity and anisotropy might need to be taken into account depending on the complexity of the model. In fact, as waveform inversion becomes more popular, more practical aspects of waveform inversion will soon become much more understood. Industrial strength software packages have recently been created (for instance GeoTomo and Wester Geco). This will certainly make it much more understood as more case studies are easily done.

Perhaps, one of the reasons the acoustic approximation is often used is that realistic viscoelastic and anisotropic modeling is computationally expensive. However, due to some theoretical developments, this trend may not continue. Source encoding is a promising method that allows lumping together the entire shot records acquired in the seismic

survey into a single encoded record. Then migration and inversion can be performed to only a single record. If this method proves robust in practice, then 3D viscoelastic waveform inversion may not be too far in the future as Virieux and Operto (2009) speculate. Waveform inversion is indeed a promising technique that would be explored much further especially as computers become even much more powerful.

This page is intentionally left blank.

Bibliography

Aki, K. and P. Richards, 2002, Quantitative seismology: University Science Press.

Al-Saleh, S. M., G. F. Margrave, and S. H. Gray, 2009, Direct downward continuation from topography using explicit wavefield extrapolation: Geophysics, **74**, S105–S112.

Al-Yahya, K., 1989, Velocity analysis by iterative profile migration: Geophysics, **54**, 718–729.

Alkhalifah, T. and C. Bagaini, 2006, Straight-rays redatuming: A fast and robust alternative to wave-equation-based datuming: Geophysics, **71**, U37–U46.

Aster, R. C., B. Borchers, and C. H. Thurber, 2005, Parameter estimation and inverse problems: Burlington, Massachusetts: Elsevier Inc.

Baysal, E., D. D. Kosloff, and J. W. C. Sherwood, 1983, Reverse time migration: Geophysics, **48**, 1514–1524.

Berkhout, A. J., 1985, Seismic migration: Elsevier Science Publ. Co., Inc., Amsterdam.

Bevc, D. and B. Biondi, 2005, Which depth imaging method should you use? a road map through the maze of possibilities: The Leading Edge, **24**, 602–606.

Billette, F., S. L. Bégat, P. Podvin, and G. Lambaré, 2003, Practical aspects and applications of 2d stereotomography: Geophysics, **68**, 1008–1021.

Billette, F. and G. Lambaré, 1998, Velocity macro-model estimation from seismic reflection data by stereotomography: Geophysical Journal International, **135**, 671–690.

Biondi, B. and P. Sava, 1999, Wave-equation migration velocity analysis: SEG Technical Program Expanded Abstracts, **18**, 1723–1726.

Biondi, B. L., 2006, 3d seismic imaging: no. 14 in Investigations in Geophysics, Society of Exploration Geophysics.

Bleistein, N., J. K. Cohen, and J. W. S. Jr., 2001, Mathematics of multidimensional seismic imaging, migration, and inversion: Springer Verlag.

Brandsberg-Dahl, S., M. V. de Hoop, and B. Ursin, 2003, Focusing in dip and ava compensation on scattering-angle/azimuth common image gathers: Geophysics, **68**, 232–254.

Brenders, A. J. and R. G. Pratt, 2007a, Efficient waveform tomography for lithospheric imaging: implications for realistic, two-dimensional acquisition geometries and low-frequency data: Geophysical Journal International, **168**, 152–170.

———, 2007b, Full waveform tomography for lithospheric imaging: results from a blind test in a realistic crustal model: Geophysical Journal International, **168**, 133–151.

Bunks, C., F. M. Saleck, S. Zaleski, and G. Chavent, 1995, Multiscale seismic waveform inversion: Geophysics, **60**, 1457–1473.

Chapman, C., 2004, Fundamentals of seismic wave propagation: Cambridge University Press.

Claerbout, J. F., 1971, Toward a unified theory of reflector mapping: Geophysics, **36**, 467–481.

———, 1985, Imaging the earth's interior: Blackwell Scientific Publications, Inc.

———, 1992, Earth soundings analysis: Processing versus inversion: Blackwell Scientific Publications, Inc.

Clayton, R. and B. Engquist, 1977, Absorbing boundary conditions for acoustic and elastic wave equations: Bulletin of the Seismological Society of America, **67**, 1529–1540.

Cox, M. J. G., 1999, static corrections for seismic reflection surveys: SEG.

Crase, E., A. Pica, M. Noble, J. McDonald, and A. Tarantola, 1990, Robust elastic nonlinear waveform inversion: Application to real data: Geophysics, **55**, 527.

Erlangga, Y. A., 2008, Advances in iterative methods and preconditioners for the helmholtz equation: Archives of Computational Methods in Engineering, **15**, 37–66.

Erlangga, Y. A., C. Vuik, and C. W. Oosterlee, 2004, On a class of preconditioners for

solving the helmholtz equation: Applied Numerical Mathematics, **50**, 409–425.

Gazdag, J., 1978, Wave equation migration with the phase-shift method: Geophysics, **43**, 1342–1351.

Gazdag, J. and P. Sguazzero, 1984, Migration of seismic data by phase shift plus interpolation: Geophysics, **49**, 124–131.

Gray, S. H., J. Etgen, J. Dellinger, and D. Whitmore, 2001, Seismic migration problems and solutions: Geophysics, **66**, 1622–1640.

Hagedoorn, J. G., 1954, A process of seismic reflection interpretation: Geophysical Prospecting, **2**, 85–127.

Jo, C.-H., C. Shin, and J. H. Suh, 1996, An optimal 9-point, finite-difference, frequency-space, 2-d scalar wave extrapolator: Geophysics, **61**, 529–537.

Lailly, P., 1983, The seismic inverse problem as a sequence of before stack migrations: in J. B. Bednar and R. Redner and E. Robinson and A. Weglein, Conference on Inverse Scattering: Theory and Application, Soc. Industr. Appl. Math., 206–220.

Lambaré, G., 2008, Stereotomography: Geophysics, **73**, VE25–VE34.

Larner, K. and L. Hatton, 1990, Wave equation migration: two approaches: First Break, **8**, 433–448.

Lavergne, M., 1961, Etude sur modele ultrasonique du probleme des couches minces en sismique refraction: Geophysical Prospecting, **9**, 60–73.

———, 1966, Refraction le long des bancs minces rapides et effet d'ecran pour les marqueurs profonds: Geophysical Prospecting, **14**, 504–527.

Liner, C., 2004, Elements of 3-d seismology: Pennwell Publishing Company Inc.

Ma, Y. and G. F. Margrave, 2008, Seismic depth imaging with the gabor transform: Geophysics, **73**, S91–S97.

Mallick, S. and L. N. Frazer, 1987, Practical aspects of reflectivity modeling: Geophysics, **52**, 1355–1364.

Margrave, G. F. and R. J. Ferguson, 1999, Wavefield extrapolation by nonstationary phase shift: Geophysics, **64**, 1067–1078.

Margrave, G. F., H. D. Geiger, S. M. Al-Saleh, and M. P. Lamoureux, 2006, Improving explicit seismic depth migration with a stabilizing wiener filter and spatial resampling: Geophysics, **71**, S111–S120.

McMechan, G. A., 1983, Migration by extrapolation of time-dependent boundary values: Geophysical Prospecting, **31**, 413–420.

Menke, W., 1989, Geophysical data analysis: Discrete inverse theory: Academic Press.

Miller, D., M. Oristaglio, and G. Beylkin, 1987, A new slant on seismic imaging: Migration and integral geometry: Geophysics, **52**, 943–964.

Mora, P., 1987, Nonlinear two-dimensional elastic inversion of multioffset seismic data: Geophysics, **52**, 1211–1228.

———, 1989, Inversion = migration + tomography: Geophysics, **54**, 1575–1586.

Mosher, C. C., S. Jin, and D. J. Foster, 2001, Migration velocity analysis using common angle image gathers: SEG Technical Program Expanded Abstracts, **20**, 889–892.

Nemeth, T., E. Normark, and F. Qin, 1997, Dynamic smoothing in crosswell traveltime tomography: Geophysics, **62**, 168–176.

Operto, S., C. Ravaut, L. Improta, J. Virieux, A. Herrero, and P. Dell'Aversana, 2004, Quantitative imaging of complex structures from dense wide-aperture seismic data by multiscale traveltime and waveform inversions: a case study: Geophysical Prospecting, **52**, 625–651.

Parker, R. L., 1994, Geophysical inverse theory: Princeton University Press.

Plessix, R.-E., 2000, Automatic cross-well tomography: an application of the differential semblance optimization to two real examples: Geophysical Prospecting, **48**, 937–951.

———, 2006, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications: Geophysical Journal International, **167**, 495–503.

———, 2009, Three-dimensional frequency-domain full-waveform inversion with an iterative solver: Geophysics, **74**, WCC149–WCC157.

Poley, J. P. and J. J. Nooteboom, 1966, Seismic refraction and screening by thin high velocity layers: Geophysical Prospecting, **14**, 184–203.

Pratt, G., C. Shin, and Hicks, 1998, Gauss-newton and full newton methods in frequency-space seismic waveform inversion: Geophysical Journal International, **133**, 341–362.

Pratt, R. G., 1990, Frequency-domain elastic wave modeling by finite differences: A tool for crosshole seismic imaging: Geophysics, **55**, 626–632.

———, 1999, Seismic waveform inversion in the frequency domain, part 1: Theory and verification in a physical scale model: Geophysics, **64**, 888–901.

———, 2008, Waveform tomography: an introduction to theory and practice: Course Notes, Queen's University, Kingston, Ontario.

Pratt, R. G., Z.-M. Song, P. Williamson, and M. Warner, 1996, Two-dimensional velocity models from wide-angle seismic data by wavefield inversion: Geophysical Journal International, **124**, 323–340.

Pratt, R. G. and M. H. Worthington, 1988, The application of diffraction tomography to cross-hole seismic data: Geophysics, **53**, 1284–1294.

Press, F., J. Oliver, and M. Ewing, 1954, Seismic model study of refractions from a layer of finite thickness: Geophysics, **19**, 388–401.

Ravaut, C., S. Operto, L. Improta, J. Virieux, A. Herrero, and P. Dell'Aversana, 2004, Multiscale imaging of complex structures from multifold wide-aperture seismic data by frequency-domain full-waveform tomography: application to a thrust belt: Geophysical Journal International, **159**, 1032–1056.

Rosenbaum, J. H., 1965, Refraction arrivals through thin high-velocity layers: Geophysics, **30**, 204–212.

Sava, P. and S. J. Hill, 2009, Overview and classification of wavefield seismic imaging

methods: The Leading Edge, **28**, 170–183.

Schneider, W. A., 1978, Integral formulation for migration in two and three dimensions: Geophysics, **43**, 49–76.

Schuster, G. T., 1998, Basics of exploration seismology and tomography, stanford mathematical geophysics summer school lectures: `http://utam.geophys.utah.edu/stanford/ch.html`. (Online; accessed 22-December-2009).

Shen, P. and W. W. Symes, 2008, Automatic velocity analysis via shot profile migration: Geophysics, **73**, VE49–VE59.

Sheng, J., A. Leeds, M. Buddensiek, and G. T. Schuster, 2006, Early arrival waveform tomography on near-surface refraction data: Geophysics, **71**, U47–U57.

Sheriff, R., 2002, Encyclopedic dictionary of applied geophysics, 4th ed: SEG.

Sheriff, R. and L. P. Geldart, 1995, Exploration seismology: Cambridge University Press.

Shragge, J., 2007, Waveform inversion by one-way wavefield extrapolation: Geophysics, **72**, A47–A50.

Sirgue, L., 2003, Inversion de la forme dónde dans le domaine frequential de donnees sismiques grands offsets.: PhD thesis, lÉcole Normale Superieure de Paris.

Sirgue, L. and R. G. Pratt, 2004a, Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies: Geophysics, **69**, 231–248.

———, 2004b, Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies: Geophysics, **69**, 231–248.

Snieder, R., 1998, The role of nonlinearity in inverse problems: Inverse Problems, **14**, 387–404.

Spencer, T. W., 1965, Refraction along a layer: Geophysics, **30**, 369–388.

Stolt, R. H. and A. K. Benson, 1986, Seismic migration: Theory and practice: Pergamon Press.

Symes, W. W., 2009, The seismic reflection inverse problem: Inverse Problems, **25**,

123008 (39pp).

Symes, W. W. and J. J. Carazzone, 1991, Velocity inversion by differential semblance optimization: Geophysics, **56**, 654–663.

Tarantola, A., 1984a, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266.

———, 1984b, Linearized inversion of seismic reflection data: Geophysical Prospecting, **32**, 998–1015.

———, 2005, Inverse problem theory and methods for model parameter estimation: SIAM.

Vermeer, G. J. O., 2002, 3-d seismic survey design: no.12 Geophysical References, Society of Exploration Geophysics.

Virieux, J. and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, **74**, WCC1–WCC26.

Štekl, I., 1997, Frequency domain seismic forward modeling: a tool for waveform inversion: PhD thesis, Department of Geology, Royal School of Mines Imperial College London.

Štekl, I. and R. G. Pratt, 1998, Accurate viscoelastic modeling by frequency-domain finite differences using rotated operators: Geophysics, **63**, 1779–1794.

Weglein, A. B., F. V. Araujo, P. M. Carvalho, R. H. Stolt, K. H. Matson, R. T. Coates, D. Corrigan, D. J. Foster, S. A. Shaw, and H. Zhang, 2003, Inverse scattering series and seismic exploration: Inverse Problems, **19**, R27–R83.

Weglein, A. B., H. Zhang, A. C. Ramírez, F. Liu, and J. E. M. Lira, 2009, Clarifying the underlying and fundamental meaning of the approximate linear inversion of seismic data: Geophysics, **74**, WCD1–WCD13.

Whitmore, N. D., 1983, Iterative depth migration by backward time propagation: SEG Technical Program Expanded Abstracts, **2**, 382–385.

Wikipedia, 2009, Geography of saudi arabia — wikipedia, the free ency-

clopedia: http://en.wikipedia.org/w/index.php?title=Geography_of_Saudi_Arabia&oldid=317815619. ([Online; accessed 9-November-2009]).

Woodward, M. J., 1992, Wave-equation tomography: Geophysics, **57**, 15–26.

Yilmaz, O., 2001, Seismic data analysis: Processing, inversion, and interpretation of seismic data: SEG.